

# Idealism

An Interpretation and Defense

Henry E. Allison

# Kant's Transcendental

REVISED & ENLARGED EDITION

# **Kant's Transcendental Idealism**

Revised and Enlarged Edition

**Henry E. Allison**

Yale University Press

New Haven and London

Copyright © 2004 by Yale University. All rights reserved.

This book may not be reproduced, in whole or in part, including illustrations, in any form (beyond that copying permitted by Sections 107 and 108 of the U.S. Copyright Law and except by reviewers for the public press), without written permission from the publishers.

Set in Adobe Garamond and Stone Sans types by The Composing Room of Michigan, Inc.

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

Allison, Henry E.

Kant's transcendental idealism / Henry E. Allison.—Rev. and enl. ed.  
p. cm.

Includes bibliographical references (p. ) and index.

ISBN 0-300-10266-6 (alk. paper)

1. Kant, Immanuel, 1724-1804. I. Title.

B2798.A634 2004

193—dc22 2003063157

A catalogue record for this book is available from the British Library.

The paper in this book meets the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council on Library Resources.

10 9 8 7 6 5 4 3 2

To Norma, with love and gratitude




# Contents

Note on Sources and Key to Abbreviations and Translations, ix  
Preface to the Revised Edition, xiii


## Part I **The Nature of Transcendental Idealism, 1**

- 1 An Introduction to the Problem, 3
- 2 Transcendental Realism and Transcendental Idealism, 20
- 3 The Thing in Itself and the Problem of Affection, 50


## Part II **Human Cognition and Its Conditions, 75**

- 4 Discursivity and Judgment, 77
- 5 The Sensible Conditions of Human Cognition, 97 
- 6 The Intellectual Conditions of Human Cognition, 133

**Part III Categories, Schemata, and Experience, 157**

- 7 The Transcendental Deduction, 159 
- 8 The Schematism of the Understanding and the Power of Judgment, 202
- 9 The Analogies of Experience, 229
- 10 Inner Sense and the Refutation of Idealism, 275

**Part IV The Transcendental Dialectic, 305**

- 11 Reason and Illusion, 307 
- 12 The Paralogisms, 333
- 13 The Antinomy of Pure Reason, 357
- 14 The Ideal of Pure Reason, 396
- 15 The Regulative Function of Reason, 423

Notes, 449

Bibliography, 515

Index, 529

## Note on Sources and Key to Abbreviations and Translations

Apart from the references to the *Critique of Pure Reason*, all references to Kant are to the volume and page of *Kants gesammelte Schriften* (KGS), herausgegeben von der Deutschen (formerly Königlichen Preussischen) Akademie der Wissenschaften, 29 vols. (Berlin: Walter de Gruyter, 1902). References to the *Critique of Pure Reason* are to the standard A and B pagination of the first and second editions, respectively. Specific works cited are referred to by means of the abbreviations listed below. The translations used are also listed below and, except in the case of the *Critique of Pure Reason*, are referred to immediately following the reference to the volume and page of the German text. It should be noted, however, that I have occasionally modified these translations. Where there is no reference to an English translation either the translation, is my own or the text is referred to but not cited.

**A/B** *Kritik der reinen Vernunft* (KGS 3–4)

*Critique of Pure Reason*, trans. and ed. by Paul Guyer and Allen W. Wood. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1997.

**Anthro** *Anthropologie in pragmatischer Hinsicht* (KGS 7)

*Anthropology from a Practical Point of View*, trans. by Mary J. Gregor. The Hague: Martinus Nijhoff, 1974.

**BL** *Logik Blomberg* (KGS 24)

*The Blomberg Logic*, in *Lectures on Logic*, trans. by Michael Young, pp. 5–246. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1992.

**Br** *Kants Briefwechsel* (KGS 10–13)

*Correspondence*, trans. by Arnulf Zweig. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1999.

**Diss** *De mundi sensibilis atque intelligibilis forma et principiis* (KGS 2)

*Concerning the Form and Principles of the Sensible and Intelligible World* (The Inaugural Dissertation), in *Theoretical Philosophy, 1755–1770*, trans. and ed. by David Walford and Ralf Meerbote, pp. 377–416. The Cambridge Edition of the Works of Immanuel Kant. Cambridge and New York: Cambridge University Press, 1992.

**EMB** *Der einzig mögliche Beweisgrund zu einer Demonstration des Dasein Gottes*

*The Only Possible Argument in Support of a Demonstration of the Existence of God*, in *Theoretical Philosophy, 1755–1770*, trans. and ed. by David Walford and Ralf Meerbote, pp. 107–202. The Cambridge Edition of the Works of Immanuel Kant. Cambridge and New York: Cambridge University Press, 1992.

**FI** *Erste Einleitung in die Kritik der Urteilkraft* (KGS 20)

“First Introduction to the Critique of Judgment”, in *Critique of the Power of Judgment*, ed. and trans. by Paul Guyer and Eric Matthews, pp. 377–416. The Cambridge Edition of the Works of Immanuel Kant. Cambridge and New York: Cambridge University Press, 2000.

**Fort** *Welches sind die wirklichen Fortschritte, die die Metaphysik seit Leibnizens und Wolf's Zeiten in Deutschland gemacht hat?* (KGS 20)

*What Real Progress Has Metaphysics Made in Germany since the Time of Leibniz and Wolff?*, trans. by Peter Heath, in *Theoretical Philosophy after 1781*, pp. 337–424. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 2002.

**Gr** *Grundlegung zur Metaphysik der Sitten* (KGS 4)

*Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, trans. by Mary J. Gregor, pp. 37–108. The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press, 2002.

**JL** *Jäsche Logik* (KGS 9)

*The Jäsche Logic*, in *Lectures on Logic*, trans. by Michael Young, pp. 521–640. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1992.

**KpV** *Kritik der praktischen Vernunft*

*Critique of Practical Reason*, in *Practical Philosophy*, trans. by Mary J. Gregor, pp. 133–272. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 2002.

**KU** *Kritik der Urteilkraft* (KGS 5)

*Critique of the Power of Judgment*, trans. by Paul Guyer and Eric Matthews. The Cambridge Edition of the Works of Immanuel Kant. Cambridge and New York: Cambridge University Press, 2000.

**LB** *Lose Blätter zur Kritik der reinen Vernunft* (KGS 23)

**LD-W** *Logik Dohna-Wundlacken* (KGS 24)

*The Dohna-Wundlacken Logic*, in *Lectures on Logic*, trans. by Michael Young, pp. 438–516. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1992.

**MAN** *Metaphysische Anfangsgründe der Naturwissenschaften* (KGS 4)

*Metaphysical Foundations of Natural Science*, in *Theoretical Philosophy after 1781*, trans. by Michael Friedman, pp. 171–270. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 2002.

**ML<sub>2</sub>** *Metaphysik L2* (KGS 28)

*Metaphysic L2*, in *Lectures on Metaphysics*, ed. and trans. by Karl Ameriks and Steve Naragon, pp. 299–356. The Cambridge Edition of the Works of Immanuel Kant. Cambridge and New York: Cambridge University Press, 2001.

**ND** *Principiorum primorum cognitionis metaphysicae nova dilucidatio* (KGS 1)

*New Elucidation*, in *Theoretical Philosophy, 1755–1770*, trans. and ed. by David Walford and Ralf Meerbote, pp. 1–46. The Cambridge Edition of the Works of Immanuel Kant. Cambridge and New York: Cambridge University Press, 1992.

**NG** *Versuch den Begriff der negativen Grössen in die Weltweisheit einzuführen* (KGS 2)

*Attempt to Introduce the Concept of Negative Magnitudes into Philosophy*, in *Theoretical Philosophy, 1755–1770*, trans. by David Walford and Ralf Meerbote, pp. 203–41. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1992.

**Pro** *Prolegomena zu einer jeden künftigen Metaphysik die als Wissenschaft wird auftreten können* (KGS 4)

*Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science*, in *Theoretical Philosophy after 1781*, trans. by Gary Hatfield, pp. 29–170. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 2002.

**PV** *Pölitz Vorlesungen über die philosophische Religionslehre* (KGS 28)

*Lectures on the philosophical doctrine of religion*, in *Religion and Rational Theology*, trans. by Allen Wood and George Di Giovanni, pp. 335–451. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1996.

**R** *Reflexionen* (KGS 15–19) and *Reflexionen Kants zur kritischen Philosophie*.

Ed. by Benno Erdmann. Leipzig: Fues's Verlag, 1882.

**Tr** *Träume eines Geistersehers, erläutert durch Träume der Metaphysik* (KGS 2)

*Dreams*, in *Theoretical Philosophy, 1755–1770*, trans. by David Walford and Ralf Meerbote, pp. 301–60. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1992.

**UE** *Über eine Entdeckung nach der alle Kritik der reinen Vernunft durch eine ältere entbehrlich gemacht werden soll* (KGS 8)

*On a Discovery whereby any New Critique of Pure Reason Is to Be Made Superfluous by an Older One*, in *Theoretical Philosophy after 1781*, trans. by Henry Allison, pp. 271–336. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 2002.

## Preface to the Revised Edition

The present work is a substantially revised version of *Kant's Transcendental Idealism*, which was first published in 1983. When I initially contemplated a new edition, I was motivated primarily by five goals: (1) to correct the more egregious errors of the initial edition; (2) to bring the first edition up to date by taking cognizance of the considerable body of secondary literature that has appeared in the past two decades; (3) to respond to my critics, where I think responses are warranted; (4) to present second thoughts concerning certain issues about which my views have evolved significantly in the intervening years; and (5) to rectify a glaring omission in the original version, namely, the neglect of the Third Analogy, for which I have been justly criticized.

Although this was already a fairly full agenda, calling for significant revision, as I began to work in earnest on the project, I gradually came to the realization that even more was required and that I would have to rethink the project in its entirety. Two factors were primarily responsible for this.

The first is the increasing tendency among Kant commentators in

recent years to argue for the separability of much of Kant's substantive theoretical philosophy from his transcendental idealism. Clearly, what I here term the "separability thesis" is not new. It was already well in place in Strawson, who endeavored to separate what he took to be the genuine analytic achievements of the *Critique* from what he regards as the "disastrous" metaphysics of transcendental idealism.<sup>1</sup> In fact, the assumption that whatever may be of philosophical value in the *Critique* must be separable from such a metaphysics has functioned virtually as an axiom in most (though not all) analytic Kant interpretations of the past century.

More recently, however, some have gone beyond this generally dismissive view of transcendental idealism, which I had already questioned in the original edition, by bringing the resources of traditional Kant scholarship, including the *Nachlass* and pre-critical writings, to bear on the issue. At least with respect to the present project, the most prominent of these Kantian "anti-idealists" are Paul Guyer and Rae Langton, whose views I consider in some detail in Chapter 1, as well as elsewhere in this book. Their especial significance for the current project stems largely from the fact that their insistence on separability is combined with a wholesale rejection of the interpretation of transcendental idealism that I presented in the first edition and subsequent writings. Assuming with Strawson that the idealism, which must be separated from the more viable insights and arguments of the *Critique*, is a disreputable form of metaphysical dogmatism, they have chided me for presenting an essentially epistemological or methodological (rather than metaphysical) interpretation of this idealism. In the terms of Guyer (endorsed by Langton), I have tried unsuccessfully to rescue Kant on this matter by trivializing his idealism, effectively reducing it to "an anodyne recommendation of epistemological modesty."<sup>2</sup>

This new situation calls for a fresh response. Specifically, it requires me not only to defend Kant's transcendental idealism and my particular interpretation of it but also to argue much more extensively and systematically than I originally did for its *inseparability* from virtually every facet of the *Critique*. Showing that transcendental idealism is a coherent, broadly defensible position is of little import unless one can also demonstrate its centrality to the Kantian project.

To this end, I develop here much more fully than I did previously the connection between transcendental idealism and what I term the "discursivity thesis," that is, the view that human cognition (as discursive) requires both concepts and (sensible) intuitions. Rather than viewing this with Strawson as the relatively unproblematic thesis that empirical knowledge rests on a duality of general concepts and particular instances, I argue both that this thesis, as Kant



understands it, marks a radical break with the epistemologies of his predecessors (rationalists and empiricists alike), and that recognizing this fact is the key to understanding his idealism. The latter is the case because it alone enables us to see how transcendental idealism is grounded in a reflection on the *a priori* conditions of human cognition (what I term “epistemic conditions”) rather than, as in other forms of idealism (for example, Berkeley’s), on the ontological status of what is known.

I believe that this increased focus on discursivity makes it possible to gain a deeper understanding of the opposition between transcendental idealism and transcendental realism, which remains the centerpiece of my interpretation. As in the first edition, everything turns on the claim that transcendental idealism and transcendental realism constitute two mutually exclusive and all-inclusive metaphilosophical alternatives or standpoints. I say “metaphilosophical” because transcendental realism, as I (and I believe Kant) understand it, encompasses a wide variety of metaphysical and epistemological views. Accordingly, what unites the various forms of such realism, many of which would not be viewed as realisms in any of the commonly accepted senses of the term, can only be an implicit commitment to a philosophical methodology, a way of analyzing metaphysical and epistemological issues that is shared by rationalist and empiricist, dogmatist and skeptic, and first challenged by Kant.

If this is true and transcendental realism stands in the disjunctive relationship to transcendental idealism suggested above, it follows that this idealism must be interpreted as itself a methodology or standpoint rather than as a substantive metaphysical doctrine. For this reason, I find it particularly noteworthy that none of my critics with whose work I am familiar (including Guyer and Langton) seriously considers the relationship between these two forms of transcendentalism. In fact, it will be argued in the body of this work that their own interpretations and criticisms (as well as those of many others) of Kant’s idealism rest on dogmatic transcendently realistic assumptions. Thus, it is hardly surprising that, from their point of view, transcendental idealism is untenable, if not unintelligible, and that the only way they can find to rescue the argument of the *Critique* is to show that it does not depend upon this idealism in a serious way. If, as I argue, transcendental idealism provides a radical challenge to the common, transcendently realistic picture of cognition, it goes without saying that it cannot be understood in terms of that picture, which is precisely what they, and Strawson before them, have attempted to do.

As in the first edition, I interpret the contrast between these two standpoints in terms of the distinction between theocentric and anthropocentric models of

cognition: transcendental realism is committed to the former, and transcendental idealism to the latter. More important, the former is equated with the intuitive and the latter with the discursive conception of cognition. Consequently, Kant's so-called Copernican revolution is to be viewed not only as a "paradigm shift" from a theocentric to an anthropocentric model but at the same time (indeed, for that very reason) also as a shift from an intuitive to a discursive conception of cognition and, therefore, a shift in our understanding of what counts as knowing. In this respect at least, Guyer and Langton are correct. Transcendental idealism, on my reading, is a doctrine of epistemological modesty, since it denies finite cognizers like ourselves any purchase on the God's-eye view of things. Nevertheless, this hardly makes it trivial or anodyne, since it also denies that the latter should be taken as the norm in terms of which human cognition is measured. On the contrary, I argue that this idealism involves a radical reconfiguration of epistemic norms, which goes together with an appeal to epistemic conditions. In other words, it serves as the epistemological counterpart of the shift from heteronomy to autonomy, which is generally recognized as the essence of Kant's "revolution" in ethics.

At first glance at least, such a claim may seem doubly paradoxical, since, when taken together with the transcendental realism–transcendental idealism dichotomy, it entails not only that the classical empiricists, in spite of their characteristic appeal to the "human understanding," were implicitly committed to the theocentric model but also that they viewed human cognition as essentially intuitive in the Kantian sense (as contrasted with discursive). Nevertheless, this is precisely what I argue in chapter 2. To anticipate, insofar as intuitive cognition is defined in opposition to the discursive variety (as it always is for Kant), it must be understood as involving an immediate and, therefore, non-conceptual grasp of its object. Since this is precisely the standard empiricist position, it bears an important structural similarity to the traditional picture of God's "way of knowing." To be sure, there is also a significant difference: whereas God's knowledge is thought to consist in a timeless, synoptic vision of the whole, most classical empiricists emphasize the narrow bounds of the human understanding, its insurmountable limitation to the materials provided to it through experience. Moreover, largely for this reason, Kant is usually thought to be on the side of the empiricists in their ongoing dispute with the rationalists, who often explicitly appeal to the theocentric picture in articulating their conception of "adequate" knowing.

Although there is undoubtedly some truth in this familiar picture of the relation between the *Critique* and classical rationalism and empiricism, it also ob-

scures a deeper point, namely, that, as forms of transcendental realism, both rationalism and empiricism share the underlying assumption that human knowledge must be measured against the norm of a putative divine cognition, which, by its very nature, must be non-conceptual. They differ mainly with respect to the question of the degree to which such intuitive cognition is attainable by the human mind. In short, from Kant's standpoint, theirs is merely a family quarrel.



The second major reason for the extent of the revisions in this edition is a basic change of mind regarding Kant's conception of reason. Although in the earlier edition I had emphasized the importance of the Antinomy of Pure Reason and discussed the B version of the Paralogisms, I had relatively little to say about Kant's general theory of reason, its connection with an underlying transcendental illusion, and its ineliminable regulative function. This was largely because I did not then fully realize the nature and depth of the connection between this theory of reason and transcendental idealism. Instead, like many I assumed that, apart from the generally discredited Fourth Paralogism (the A version's refutation of idealism), this idealism figured prominently only in the Antinomy, where it supposedly provided an escape from the contradictions in which reason found itself entangled on transcendentially realistic principles. Moreover, this in turn led naturally to the assumption that the transcendental illusion to which reason was subject is itself a product of transcendental realism and, therefore, will itself disappear with the rejection of the latter.



I was awakened from my "dogmatic slumber" on this issue, however, by the work of a former student, Michelle Grier. First in her dissertation and then, more substantively, in an important book based upon it, Grier has shown conclusively that for Kant transcendental illusion is inherent in the very nature of human reason.<sup>3</sup> For present purposes, the importance of this fact, which is sometimes noted but seldom taken seriously, is that it shows the necessity of drawing a sharp distinction between transcendental illusion and transcendental realism. Since the former is inherent in the theoretical use of reason, it cannot be eliminated (though the metaphysical errors stemming from it can be avoided), whereas transcendental realism, as a metaphilosophical stance, can (and should) be replaced by transcendental idealism. But if this is true, it follows that transcendental idealism does not, as I had previously assumed, eliminate transcendental illusion (that being an impossible task). Its service is rather to prevent us from being deceived by this illusion, which it accomplishes by separating it from the transcendental realism with which it is commonly conjoined. The latter thus remains the real source of the difficulties in which reason

finds itself when, under the spell of transcendental illusion, it ventures into the transcendent.

This gives to transcendental idealism an essential therapeutic function of which I had not initially been fully aware, but which accords with my basic understanding of this idealism and of its inseparability from the overall argument of the *Critique* (its critical as well as its constructive parts). In order to develop this point, however, I have found it necessary to provide a systematic interpretation of Kant's theory of reason. In addition to addressing some of the familiar criticisms of this theory, this requires discussing the general account of reason and its connection with illusion, which is spelled out in the Introduction to the Dialectic, as well as its role within the Paralogisms and the Ideal, and, finally, its regulative function as articulated in the Appendix to the Dialectic. This accounts for the presence of three completely new chapters (11, 14, and 15), as well as the relocation and significant modification of the chapters dealing with the Paralogisms and the Antinomy (12 and 13 in the present version). In short, it has led to a total recasting of the fourth part of this work.

For the reader interested in comparing the present with the previous version of this work, four other changes are worthy of note. The first, which is of a more substantive nature, concerns the treatment of the Metaphysical Deduction (chapter 6). Like most commentators, I had previously thought that there was not much to be said about the specifics of Kant's attempt to establish the completeness of his table of the functions of judgment from which the table of categories is supposedly derived. Instead, I downplayed the significance of the completeness thesis and focused on the connection between certain judgmental functions and the corresponding categories, which is an important and deeply controversial issue in its own right. Once again, however, I was led to change my mind by the illuminating discussions of this issue that have appeared in recent years.<sup>4</sup> As a result of this recent work, I have greatly expanded my discussion in chapter 6, arguing not merely for the systematic importance of the completeness question but also for the plausibility of Kant's position, given the discursivity thesis with which it is closely connected. Whether this case for completeness is convincing is left for the reader to decide.

The other notable changes are largely of a structural nature. The first of these is a direct consequence of the inclusion of a discussion of the Third Analogy. If this discussion is to be anything more than perfunctory (as treatments of the Third Analogy often are), it calls for either the addition of a new chapter dealing with this Analogy or a major reorganization of the discussion of the Analogies as a whole. Although it results in a somewhat unwieldy chapter, I have

opted for the second course, combining the discussion of all three Analogies into a single chapter (9). The main motivation for this is that it makes it possible to treat the important theme of the interconnection of these Analogies, a topic I had largely neglected in the original edition.

Another significant structural change concerns the Refutation of Idealism. Whereas I had previously devoted separate chapters to the Refutation and to Kant's theory of inner sense, they are here combined in a single chapter (10), a procedure I believe to be justified by the intimate connection between these topics. In this context, I would also like to call attention to the fact that I have added a discussion of the connection of the Refutation of Idealism and transcendental idealism, which was lacking in the original version. Usually, the suggestion of such a connection is greeted with deep suspicion, the assumption being that any refutation of idealism that rests on the latter cannot be a genuine refutation. But if, as I maintain, transcendental idealism is inseparable from the substantive doctrines of the *Critique*, it should also play a significant role here, in perhaps the least likely place. Thus, in opposition to Guyer and many others, I attempt to argue that this idealism (correctly understood) does, in fact, play such a role in the B version's Refutation of Idealism. Indeed, I suggest that it is an essential pre-condition of *any* successful refutation of what Kant sometimes terms the "common idealism," and that the refusal to acknowledge this reflects a deep-seated misunderstanding of transcendental idealism.

Finally, there is a significant omission in this revised version of *Kant's Transcendental Idealism*. It concerns the discussion of the problem of freedom, which in many ways lies at the very heart of Kant's philosophy. To be sure, the topic is not neglected entirely, since I do deal with what may be called the cosmological dimension of the issue in connection with the discussion of the Third Antinomy (chapter 13). Unlike in the original version, however, I do not attempt to connect this to the question of free will (practical freedom), where most of the interest lies. Although regrettable, this omission was necessitated by the overriding need to keep this work within a manageable size. Moreover, those who may be interested in learning my views on the matter can consult my fuller treatment of it elsewhere.<sup>5</sup>

## Acknowledgments

In addition to all those whose assistance I acknowledged in the original edition of this book, I would like to thank those to whom I am indebted for their help with this revised edition. Foremost among these is my former student, Michelle Grier, not only for her own dissertation and book, which forced me to rethink a number of my earlier views, but also for her insightful comments and suggestions concerning drafts of the chapters dealing with Kant's theory of reason. I am also indebted to another former student, Luigi Caranti, who wrote an excellent dissertation on Kant's refutation of idealism, and with whom I have had many discussions on this topic in recent years. I would like to express my appreciation to Katrien Vander Straeten for her very material assistance in the preparation of this manuscript and the index. Finally, I would like to thank Cambridge University Press for permission to quote extensively from the translation of the *Critique of Pure Reason* by Paul Guyer and Allen Wood.


**Part One The Nature of  
Transcendental Idealism**


## Chapter 1 An Introduction to the Problem

In spite of some sympathy shown in recent years for a vaguely Kantian sort of idealism or, better, anti-realism, which argues for the dependence of our conception of reality on our concepts and/or linguistic practices, Kant's transcendental idealism proper, with its distinction between appearances and things in themselves, remains highly unpopular.<sup>1</sup> To be sure, there has arisen a lively dispute concerning the interpretation of this idealism, with some, myself included, arguing for a version of what is usually called a "two-aspect" view (to be discussed below). Nevertheless, many interpreters continue to attribute to Kant the traditional "two-object" or "two-world" view or some close facsimile thereof, and in most (though not all) cases this reading is combined with a summary dismissal of transcendental idealism as a viable philosophical position. In fact, the manifest untenability of transcendental idealism, as they understand it, has led some critics to attempt to save Kant from himself, by separating what they take to be a legitimate core of Kantian argument (usually of an anti-skeptical nature) from the excess baggage of transcendental idealism, with which they believe it to be encumbered.





The approach taken in the present work is diametrically opposed to this. Although not denying many of the difficulties pointed out by the critics, its main goal is to provide an overall interpretation and, where possible, a defense of transcendental idealism. This defense will not amount to an attempt to demonstrate the truth of transcendental idealism; that being much too ambitious a project. It will, however, argue that this idealism remains a viable philosophical option, still worthy of serious consideration. An underlying thesis, which is independent of the viability of transcendental idealism, is its intimate connection with virtually every aspect of the *Critique*. In short, the separability of Kant's fundamental claims in the *Critique* from transcendental idealism will be categorically denied. For better or worse, they stand or fall together. 

The present chapter is intended as an introduction to the problem as a whole and is divided into two main parts. The first provides a brief sketch of what might be termed the “anti-idealist” reading of Kant. The second begins the process of rehabilitation by outlining a conception of transcendental idealism quite different from the one dismissed by its critics. It argues that this idealism is more properly seen as epistemological or perhaps “metaepistemological” than as metaphysical in nature, since it is grounded in an analysis of the discursive nature of human cognition.<sup>2</sup> To this end, it introduces the concept of an “epistemic condition” as a key to the understanding not only of transcendental idealism but also of the argument of the *Critique* as a whole. 

## I. KANTIAN ANTI-IDEALISM

As noted above, critics of transcendental idealism, who are nonetheless in some sense sympathetic to Kant, tend to interpret this idealism in an extremely uncharitable manner and then argue for its separability from some independently justifiable strand of Kantian argumentation. We shall briefly consider each of these moves in turn.

### A. The Idealism of Anti-Idealism

According to many of its critics, transcendental idealism is a metaphysical theory that affirms the uncognizability of the “real” (things in themselves) and relegates cognition to the purely subjective realm of representations (appearances). It thus combines a phenomenalist, essentially Berkeleian, account of what is actually experienced by the mind, and therefore cognizable, namely, its own representations, with the postulation of an additional set of entities, which, in terms of the very theory, are uncognizable.<sup>3</sup> In spite of the obvious

paradox it involves, this postulation is deemed necessary to explain how the mind acquires its representations, or at least the materials for them (their form being “imposed” by the mind itself). The basic assumption is simply that the mind can acquire these materials only as a result of being “affected” by things in themselves. Thus, such things must be assumed to exist, even though the theory denies that we have the right to say anything about them, including the claims that they exist and affect us.

Although it has a long and reasonably distinguished lineage, traceable to Kant’s contemporaries,<sup>4</sup> the continued acceptance of this understanding of transcendental idealism in the Anglo-American philosophical community is largely due to the influence of P. F. Strawson, who brusquely defines this idealism as the doctrine that “reality is supersensible and that we can have no knowledge of it.”<sup>5</sup> Moreover, in the spirit of this reading, Strawson not only rejects transcendental idealism as incoherent; he also provides an account of what led Kant to this “disastrous” doctrine. As Strawson sees it, transcendental idealism is the direct consequence of Kant’s “perversion” of the “scientifically minded philosopher’s” contrast between a realm of physical objects composed of primary qualities and a mental realm consisting of the sensible appearances of these objects (including their secondary qualities). This mental realm, like its Kantian counterpart, is thought to be produced by means of an affection of the mind by physical objects. Kant allegedly perverts this model, however, by assigning the whole spatiotemporal framework (which according to the original model pertains to the “real,” that is to say, to physical objects) to the subjective constitution of the human mind. The resulting doctrine is judged to be incoherent because, among other reasons, it is only with reference to a spatiotemporal framework that one can talk intelligibly about “affection.”<sup>6</sup>



In addition to its unwarranted postulation of things in themselves that somehow affect the mind, transcendental idealism is often attacked on epistemological grounds for its complementary claim that we can know only appearances. Equating Kantian “appearances” with “mere representations,” critics take this to mean that we know only the contents of our own minds, that is, ideas in the Berkeleian sense. This is then sometimes used as the basis for a critique of the doctrine of the ideality of space and time, which Kant presents in the *Transcendental Aesthetic*. Simply put, the claim is that Kant’s subjectivistic starting point confronts him with a stark dilemma: he must maintain either (1) that things only seem to us to be spatial, or (2) that appearances, that is to say, representations, really are spatial. The former, however, allegedly entails that our consciousness of a world of objects extended and located in space is some-

how illusory; whereas the latter is supposedly absurd on the face of it, since it requires us to regard mental items as extended and located in space.<sup>7</sup>

This line of criticism can likewise be traced back to Kant's contemporaries, and it certainly has echoes in Strawson.<sup>8</sup> Perhaps its sharpest twentieth-century formulation, however, is by H. A. Prichard, who concentrates much of his attack on the alleged incoherence of Kantian "appearance talk." According to Prichard, whose critique was highly influential in the earlier part of the century, Kant's whole conception of appearance is vitiated by a confusion of the claim that we know only things as they appear to us with the quite different claim that we know only a particular class of things, namely, appearances. Prichard also suggests that Kant only managed to avoid the above-mentioned dilemma by sliding from one of these claims to the other. Thus, on his reconstruction, what Kant really wished to hold is that we know things only as they appear to us; but since this supposedly entails that these things only *seem* to us to be spatial, in order to defend his cherished empirical realism, he was forced to shift to the doctrine that we know appearances and that they really are spatial.<sup>9</sup>

Underlying Prichard's critique is the assumption that the claim, that we know objects only as they appear, is to be understood to mean that we know only how they seem to us, not how they really are. In fact, he makes this quite explicit by construing Kant's distinction in terms of the classic example of perceptual illusion: the straight stick that appears bent to an observer when it is immersed in water. And, given this, he has little difficulty in reducing Kant's doctrine to absurdity. Although his analysis proceeds through various stages, Prichard's main point is linguistic. Specifically, he claims that Kant's account contradicts the clear meaning of 'knowledge'. Since to know something, according to Prichard, just means to know it as it really is, in contrast to how it may "seem to us," it follows that for Kant we cannot really know anything at all.<sup>10</sup> Thus, far from providing an antidote to skepticism, as was his intent, Kant, on this reading, is seen as a Cartesian skeptic *malgré lui*.

## B. The Separability Thesis

What is here termed the separability thesis is a response to this understanding of transcendental idealism and is likewise largely the work of P. F. Strawson, who set as his avowed task the separation of what he terms the "analytic argument" of the *Critique* from the transcendental idealism with which he believes Kant unfortunately and unnecessarily entangled it.<sup>11</sup> Central to the former, as Strawson conceives it, is the refutation of a Cartesian-type skepticism through the demonstration of a connection between self-consciousness (or the self-

ascription of experience) and the experience of a public, objective world. In this respect Strawson has been followed by numerous philosophers, who have tried to formulate and defend some vaguely Kantian anti-skeptical or “transcendental” arguments that are uncontaminated by any idealistic premises.<sup>12</sup>

Although for Strawson himself the project is more one of appropriating what is deemed valuable in Kant and discarding the rest than of proposing a radically new interpretation based on a careful consideration of the relevant texts, the latter has been attempted by two Kant scholars who appear to have been deeply influenced by Strawson’s work: Paul Guyer and Rae Langton. Accordingly, it may prove instructive to examine their views on transcendental idealism and the separability thesis at this point.

*Guyer.* In dismissing interpretations of transcendental idealism such as the one presented in the first edition of this book, which he characterizes as an “anodyne recommendation of epistemological modesty,”<sup>13</sup> Guyer insists on the dogmatic metaphysical character of this idealism. As he puts it, “Transcendental idealism is not a skeptical reminder that we *cannot be sure* that things as they are in themselves *are* also as we represent them to be; it is a harshly dogmatic insistence that we *can be quite sure* that things as they are in themselves *cannot be* as we represent them to be.”<sup>14</sup> And since space and time are the indispensable elements in our representations of things, he proceeds to identify this idealism with the thesis that “things in themselves, whatever else they may be, *are not* spatial and temporal.”<sup>15</sup>

Although we shall see that Guyer is correct to insist that Kant affirmed the strong thesis of the non-spatiotemporality of things as they are in themselves rather than the weaker thesis that we cannot be sure about the matter, this need not make Kant’s position “harshly dogmatic,” or even dogmatic at all, for that matter. Moreover, since Kant explicitly denies that we *represent* things as they are in themselves as spatial or temporal, Guyer’s claim that for Kant things as they are in themselves are not as we represent them to be is somewhat puzzling.

It is, however, possible to make some sense of this once one understands how Guyer views the concept of the thing in itself. Interestingly enough, this emerges from his dismissive treatment of the two-aspect view (to be discussed below). Although Guyer acknowledges that, except in the special cases of God and the soul, Kant does not “postulate a second set of ghostlike nonspatial and nontemporal objects in addition to the ordinary referents of empirical judgments,” he insists that this is of no avail to defenders of transcendental idealism, since “he does something just as unpleasant—namely, *degrade* ordinary objects

to mere representations of themselves, or identify objects possessing spatial and temporal properties with mere mental entities.”<sup>16</sup> Moreover, Guyer suggests that Kant has no need to postulate a distinct set of objects underlying appearances in order to affirm the non-spatial and non-temporal nature of things in themselves, “because the ontology from which he begins *already* includes two classes of objects, namely things like tables and chairs and our representations of them.”<sup>17</sup>

As Guyer proceeds to make clear, his view is that Kant’s idealism turns on (or consists in) the transference of spatial and temporal properties from the ordinary objects of human experience to appearances, understood as mere representations or mental entities.<sup>18</sup> In short, he accepts the gist of Strawson’s account of how Kant was (mis)led to this disastrous doctrine. Consequently, he thinks that Kant has no need to posit an additional set of objects that are not in space or time. But this suggests that by things in themselves Guyer must understand the ordinary objects of experience, such as tables and chairs, stripped of their spatial and temporal properties. And this is presumably why he claims that for Kant we can know with certainty that things in themselves are *not* as we represent them as being, that is, as in space and time.

Although this conception of the thing itself may seem to be suggested by certain passages in which Kant attempts to illustrate the ideality of appearances by extending the subjectivity of secondary qualities to include the primary ones as well (all of which involve some reference to space and the dynamical conditions of filling it),<sup>19</sup> the conception borders on incoherence if it is taken as anything more than a loose analogy. For it requires us to read Kant as *both* identifying the ordinary objects of human experience with things in themselves *and* denying that these things possess the properties that we supposedly experience them as having.

Guyer is unperturbed by any such incoherence, however, because he does not think it matters. At least he does not think it matters to what he terms Kant’s “transcendental theory of experience,” that is, the claims of the Analogies and especially the Refutation of Idealism, which is all that he finds worth preserving in the *Critique*. Thus, in response to Jacobi’s famous remark that “*without* the presupposition [of the thing in itself] I cannot enter the system, and with that presupposition I cannot remain in it”<sup>20</sup> (to which we shall return in chapter 3), Guyer comments:

One can enter the critical philosophy, or at least the transcendental theory of experience, without the presupposition of the thing in itself, because none of Kant’s argu-



ments for the nonspatiality and nontemporality of things in themselves, certainly none of his arguments from legitimate claims of the transcendental theory of experience, succeeds. Thus one can accept the transcendental theory of experience finally expounded in the analogies of experience and the refutation of idealism without any commitment to dogmatic transcendental idealism.<sup>21</sup>



Whether Kant's arguments for idealism fail, as Guyer suggests, can be decided only by their careful consideration. But it also remains to be seen whether Kant's transcendental theory of experience is really separable from this idealism, when the latter is interpreted in a more sympathetic manner than Guyer is prepared to countenance.<sup>22</sup> As already noted, one of the major concerns of this work is to show that it is not.

*Langton.* Whereas Guyer concedes that there is an idealistic strand in Kant's thought, which he believes to be separable from its defensible core, Rae Langton depicts a Kant who is in all essential respects a robust realist, indeed, a scientific realist in the contemporary sense, suitably equipped with a causal theory of knowledge.<sup>23</sup> Naturally, she does not deny that there are some passages that are not readily amenable to such a reading, but she endeavors to minimize them and to argue that they do not commit Kant to idealism in any significant way.<sup>24</sup> Nevertheless, in sharp contrast to Guyer, she is not dismissive of Kant's distinction between appearances and things in themselves and the denial of the possibility of cognizing the latter. On the contrary, she insists that Kant's "humility" about the cognition of things in themselves is not only compatible with, but actually required by, his realism. Accordingly, in spite of significant differences, Langton is at one with Guyer in rejecting any "anodyne" reading or, as she terms it, "deflationary proposal," regarding such humility.<sup>25</sup>

The starting point of Langton's unabashedly realistic reconstruction of Kant's contrast between appearances and things in themselves, and of the limitation of knowledge to the former, is Strawson's exegetical thesis that humility supposedly follows from receptivity, that is, that the reason things in themselves are unknowable by the human mind is that our cognition is receptive and, therefore, dependent on being affected by the object cognized.<sup>26</sup> But whereas Strawson treats this as a fundamental and unargued premise of the *Critique*, Langton, recognizing that humility does not follow directly from receptivity, acknowledges the need for an additional premise linking the two. In fact, the major aim of her book appears to be to provide such a premise.

This need would be readily acknowledged by more orthodox Kantians, who

tend to locate the required premise in Kant's conception of human sensibility and of space and time as forms thereof. Langton denies, however, that Kant's humility regarding the cognition of things in themselves has anything to do with his views about space and time and their connection with human sensibility.<sup>27</sup> Instead, she locates the missing premise in an anti-Leibnizian metaphysical thesis concerning the irreducibility of relations, which is traceable to some of Kant's earliest writings and which supposedly survives in the *Critique*.

Reduced to its essentials, Langton's reconstruction of Kant's humility argument consists of three steps. The first is the characterization of things in themselves as substances having intrinsic properties and phenomena as relational properties of these substances. The second is the claim that the relations and relational properties of such substances are not reducible to their intrinsic properties. In the more contemporary terms Langton prefers, this means that the relational properties of things do not supervene on their intrinsic properties. The third and final step is just the appeal to receptivity emphasized by Strawson. Since the properties through which things affect us (their causal powers) are merely relational and, as such, do not supervene on their intrinsic properties, and since we must be affected by an object in order to cognize it, it follows that we cannot know its intrinsic properties, which is just the doctrine of humility.<sup>28</sup>

An essential feature of Langton's account is the virtual identification of Kantian things in themselves with Leibnizian monads (substances with intrinsic properties). Moreover, this is certainly a controversial thesis, to say the least. For though, as we shall see, the "critical" Kant has good reasons for characterizing the *concept* of a thing as it is in itself in such a manner, it hardly follows that Kant remained committed to the substantive metaphysical view that reality is composed of such substances. Indeed, this invites the obvious question: How is such a thesis, which is necessary to ground Kantian humility on Langton's reading, compatible with this very humility?

Langton is well aware of this objection, and she attempts to deflect it by appealing to what she takes to be required for the cognition of a thing as it is in itself (if we could have it). In addition she assigns a cognitive function to the pure (as opposed to the schematized) concept of substance. Such cognition, she suggests, would require a capacity to determine a substance by ascribing intrinsic predicates to it, which the human mind is incapable of doing in virtue of the receptive nature of its cognition. Nevertheless, she maintains that this is perfectly consonant with the application to such a thing of the pure concept of substance. As she puts it:

It is compatible with this [the impossibility of knowledge of its intrinsic properties] that one can use the pure concept in a manner which will allow one to assert the existence of substances, and to assert that they must have intrinsic properties: for this use falls short of a use that attempts to determine a thing by ascribing to it particular distinctive and intrinsic predicates.<sup>29</sup>

Although Langton is certainly correct to point out that such a use falls short of the attempt to determine what intrinsic properties a substance possesses, this does not suffice to legitimize the minimal, yet nonetheless metaphysical, use that she wants to allow.<sup>30</sup> In fact, we shall see that Kant explicitly disallows any such use on the grounds that it involves what he terms a “transcendental mis-employment” of the categories. But in order to appreciate this, we must first understand the function of the categories as epistemic conditions, which, in turn, rests on a prior understanding of the general concept of such a condition as it relates to discursive cognition.

## II. EPISTEMIC CONDITIONS, DISCURSIVITY, AND TRANSCENDENTAL IDEALITY

By an epistemic condition is here understood a necessary condition for the representation of objects, that is, a condition without which our representations would not relate to objects or, equivalently, possess objective reality. As such, it could also be termed an “objectivating condition,” since it fulfills an objectivating function. As conditions of the possibility of representing objects, epistemic conditions are distinguished from both psychological and ontological conditions. By the former is meant a propensity or mechanism of the mind, which governs belief and belief acquisition. Hume’s custom or habit is a prime example of such a condition. By the latter is meant a condition of the possibility of the existence of things, which conditions these things quite independently of their relation to the human (or any other) mind. Newton’s absolute space and time are conditions in this sense. Epistemic conditions share with the former the property of being “subjective,” that is, they reflect the structure and operations of the human mind. They differ from them with respect to their objectivating function. Correlatively, they share with the latter the property of being objective or objectivating. They differ in that they condition the objectivity of our *representations* of things rather than the very existence of the things themselves. As we shall see, the fundamental problem confronting transcendental idealism is to explain how such conditions can be both subjective and objective or objectivating at once.






This emphasis on an objectivating function is crucial, since not everything that one could regard as a condition of cognition counts as an epistemic condition in the intended sense. Thus, critics intent on denying any link between conditions of cognition and Kant's idealism point to empirical examples, such as the fact that our eyes can perceive things only if they reflect light of a certain wavelength. As a fact about our visual capacities, this is arguably a "condition" of a significant subset of the perceptual cognition of sighted human beings; but this hardly brings with it any idealistic implications. And the same may be said about the other sensory modalities, each of which involves built-in restrictions on the range of data that can be received and processed.<sup>31</sup>

All of this is certainly true, but beside the point. Conditions of this sort are not epistemic in the relevant sense, because they have no objective validity or objectivating function. On the contrary, as with the Humean psychological conditions, an appeal to them presupposes the existence of an objective, spatiotemporal world, the representation of which is supposedly to be explained. Accordingly, it hardly follows from the fact that such conditions do not entail any sort of idealism, that properly epistemic conditions, if there are such, do not do so either.

In fact, the concept of an epistemic condition brings with it an idealistic commitment of at least the indeterminate sort noted at the beginning of this chapter, because it involves the relativization of the *concept* of an object to human cognition and the conditions of its representation of objects. In other words, the claim is not that *things* transcending the conditions of human cognition cannot exist (this would make these conditions ontological rather than epistemic) but merely that such things cannot count as *objects* for us. This also appears to be the sense of Kant's famous "Copernican hypothesis" that objects must "conform to our cognition" (*sich nach unserem Erkenntniss richten*) (Bxvi). As we shall see in chapter 2, this means that objects must conform to the conditions of their representation; not that they exist in the mind in the manner of Berkeleian ideas or the sense data of phenomenologists.


Nevertheless, this broad concept of an epistemic condition is not sufficient to capture what is distinctive in Kant's transcendental idealism.<sup>32</sup> The latter does not merely relativize the concept of an object to the conditions (whatever they may be) of the representation of objects, it also specifies these conditions by means of an analysis of the discursive nature of human cognition. Consequently, Kant's idealism depends crucially on his conception of human cognition as discursive, what we shall henceforth call the discursivity thesis.

Since this thesis will be discussed in some detail in subsequent chapters, it

must here suffice to note that to claim that human cognition is discursive is to claim that it requires both concepts and *sensible* intuition. Without the former there would be no thought and, therefore, no cognition; without the latter there would be nothing to be thought. As Kant puts it in an oft-cited phrase, “Thoughts without content are empty, intuitions without concepts are blind” (A51 / B76). 

Although some of Kant’s idealistic successors challenged this thesis, at least insofar as it concerns philosophical knowledge,<sup>33</sup> the reaction of many modern-day analytical commentators, such as Strawson, is quite different. Indeed, for Strawson, the discursivity thesis reduces to the inescapable necessity in any philosophical thinking about experience or empirical knowledge to assume a “duality of general concepts, on the one hand, and particular instances of general concepts, encountered in experience, on the other.”<sup>34</sup> Accordingly, for him the problem lies in the fact that Kant did not rest content with this plain truth; instead, interpreting it psychologically (or at least expressing it in a psychological idiom), he linked each condition to a distinct cognitive faculty. But with this he was led to that “disastrous model” in which the mind is viewed as imposing its forms on things.<sup>35</sup>

We shall see, however, that this dismissive treatment ignores certain essential features of Kant’s account. In particular, it ignores the fact (to be explored in chapter 2) that since the discursivity thesis was denied (at least as an account of adequate cognition) by Kant’s predecessors, both rationalist and empiricist, it cannot be the innocuous, non-controversial thesis that Strawson takes it to be.<sup>36</sup>

Admittedly, Kant himself is not completely innocent on this score. At least part of the problem is that he tends to argue from rather than for the discursivity thesis, thereby suggesting that he viewed it as an unquestioned presupposition or starting point rather than as something that itself stands in need of justification. Nevertheless, at least the outline of an argument for this thesis is implicit in the *Critique*.<sup>37</sup> 

The underlying assumption of this argument, which can be sketched here only in the baldest terms, is that cognition requires that an object somehow be given to the mind. In Kant’s terminology, this means that it must be present (or presentable) in intuition, by which he understands a singular representation that is immediately related to its object (A320 / B377).<sup>38</sup> Although Kant never says so explicitly, this appears to be a general claim, applicable to both divine or intuitive and human or discursive cognition. Kant further assumes that there are only two conceivable types of intuition: sensible and non-sensible or intellectual. But since the latter, as Kant conceives of it, requires the actual genera-

tion of the object through the act of intuiting, that is, a creative intuition, it is ruled out for human cognizers as incompatible with our finitude.<sup>39</sup>

It follows from this that our intuition and, more generally, that of any finite cognizer, must be sensible, that is, receptive, resulting from an affection of the mind by objects. This is still not sufficient to establish the discursivity thesis, however, since the latter explicitly affirms the active, *conceptual* nature of cognition, not merely its dependence on receptivity. In fact, empiricists like Berkeley and Hume, with their basically imagistic account of thinking, would readily grant receptivity, while denying discursivity. Consequently, the full argument for this thesis requires the additional premise that sensible intuition alone is not *sufficient* to yield cognition of objects, that it provides the data for such cognition but does not itself amount to cognition.

Since this anti-empiricist premise regarding the constitutive role of the understanding is the central claim of the Transcendental Analytic, it cannot be examined here. Nevertheless, it is important to note that it carries with it a twofold lesson for the understanding of Kant's theory of sensibility. On the one hand, as we have just seen, it entails that sensibility is merely receptive and, therefore, capable only of providing the raw data for cognition (otherwise the spontaneity of the understanding would not be required). On the other hand, and somewhat surprisingly, it also means that the very possibility of discursive cognition requires that the data be presented by sensibility in a manner suitable for conceptualization. Accordingly, even though sensibility does not itself order the given data, that being the task of the understanding, or, more properly, the imagination, it must present them in such a way that they are "capable of being ordered" (A29 / B34).<sup>40</sup> And the latter, it will be argued, provides the basis for Kant's idealism.

The crucial point is that for Kant this original orderability, as well as the actual ordering, is a contribution of the cognitive subject and that this marks his decisive break with empiricistic (as well as rationalistic) accounts of sensibility. Kant already hints at this in his initial account of receptivity in the *Transcendental Aesthetic*, where he defines it or sensibility (he equates the two) as "the capacity [*Fähigkeit*] to acquire representations through the way [*die Art*] in which we are affected by objects" (A19 / B33). As this definition clearly indicates, sensibility involves not merely a capacity for being affected by objects, and, therefore, for receiving sensory data, but also for being affected in a "certain way" or "manner" (*Art*). This means that the way in which sensibility presents its data to the understanding for its conceptualization already reflects a particular manner of receiving it, that is, a certain form of sensibly intuiting,

which is determined by the nature of human sensibility rather than by the affecting objects. Moreover, as we shall see later, this form of sensibly intuiting conditions the possibility of its ordering by the understanding.

Admittedly, the connection of this account of sensibility with the discursivity thesis is not immediately apparent. For it might very well seem that the spontaneity of the understanding could operate on raw sensible data, which are unencumbered by any *a priori* forms or conditions. Nevertheless, a closer consideration suggests that this is not the case, at least not if we understand discursivity in the Kantian sense, as requiring the joint contribution of sensibility and understanding.

There appear to be two possible ways in which such a scenario might be understood, neither of which yields a viable account of discursivity. One is to view sensibility as presenting to the understanding objects as they are in themselves (not as they appear in virtue of subjective conditions of sensibility). This is the traditional pre-Kantian view, and the problem is that under this scenario thought would have to be viewed either with Leibniz as exercising merely a clarificatory function (bringing conceptual clarity and distinctness to what the senses present obscurely), with Locke as creating its own “nominal essences” (the “workmanship of the understanding”), which are of pragmatic value but do not provide genuine cognition, or with Hume as copying lively impressions in the dimmer medium of ideas. But in none of these cases is there room for any genuine spontaneity or, as Kant sometimes put it, a “real use of the understanding.”

Alternatively, in order to find room for the latter, one might assume that what is given “in itself” are not objects but the data for the thought of objects that must still be unified by the understanding to yield full-fledged cognition. In short, we would have the Transcendental Analytic without the Transcendental Aesthetic (a prospect that might seem attractive to many Kant interpreters, including Strawson and Guyer). But, in spite of its superficial attractiveness, it seems clear that Kant would not find it appealing. For the only sense that could be made of the idea that the sensible data for the thought of objects (as contrasted with the objects thought) present themselves as they are in themselves is that the successive temporal order of their appearing is one that pertains to them as they are in themselves (independently of their relation to human sensibility). In that case, however, it is difficult to see how thought could gain any purchase on the world or any claim to objectivity. To anticipate the argument of the Analogies, there would be no room for an objective ordering of states and events as distinct from a subjective ordering of perceptions. Either the two or-

ders would simply coincide, which amounts to phenomenalism, or there would be no way, short of metaphysical assumptions, such as a pre-established harmony, for getting from the one to the other, which leads either to skepticism or to an ungrounded dogmatism. Thus, paradoxically enough, it is precisely the *denial* that sensibility makes an autonomous *a priori* contribution to cognition that entails these unattractive options.

Like much else in this preliminary discussion, the above is a substantive and controversial thesis that will be further explored in the body of this work. What must be emphasized at present is simply that, assuming the correctness of the preceding analysis, we have the basis for an understanding of transcendental idealism that is rooted in Kantian epistemology rather than a pre-critical metaphysics, and that is for this reason far more philosophically attractive than the dismissive criticisms of it lead one to believe.

As I shall argue at greater length, this epistemologically based understanding of transcendental idealism requires that the transcendental distinction between appearances and things in themselves be understood as holding between two ways of *considering* things (as they appear and as they are in themselves) rather than as, on the more traditional reading, between two ontologically distinct sets of entities (appearances and things in themselves). In this regard, it may be characterized as a “two-aspect” reading. Nevertheless, this label requires careful qualification in order to avoid serious misunderstanding. The basic problem is that dual- (or multi-) aspect theories are themselves usually metaphysical in nature. In fact, they typically arise in connection with treatments of the mind-body problem, where some version of “dual aspectism” is sometimes proposed as a viable alternative to both dualism and materialism.<sup>41</sup> The classical example of such a metaphysical dual-aspect theory is Spinoza’s account of the mind as constituting one and the same thing as the body.<sup>42</sup> Perhaps the best known contemporary version of it is Davidson’s “anomalous monism,” which because of its assertion of a token-token identity between physical and mental states has been suggested as a model for interpreting Kant’s transcendental idealism.<sup>43</sup>

The main problem with attempting to interpret transcendental idealism on the basis of such a metaphysical model is that it loses sight of its fundamentally epistemological thrust, which is itself the result of approaching it independently of the discursivity thesis. As was argued above, this thesis entails that sensibility must have *some a priori* forms (though not that they be space and time). Accordingly, in considering things as they appear, we are considering them in the way in which they are presented to discursive knowers with our forms of sensibility. Conversely, to consider them as they are in themselves is to

consider them apart from their epistemic relation to these forms or epistemic conditions, which, if it is to have any content, must be equivalent to considering them qua objects for some pure intelligence or “mere understanding.”<sup>44</sup> It is the qualitative (transcendental) distinction between the sensible and the intellectual conditions of discursive cognition that makes this dual manner of consideration possible, just as it is the dependence of thought on sensibility for its content that prevents the latter mode of consideration from amounting to cognition.



When Kant's distinction is understood in this way, the claim that we can cognize things only as they appear, not as they are themselves, need not be taken (as it was, for example, by Prichard) to mean that we can know only how things seem to us under certain conditions or through a “veil of perception.” Rather, such cognition is fully objective, since it is governed by *a priori* epistemic conditions. It is only that, as discursive, human knowledge differs in kind, not merely in degree, from that which might be had by a putative pure understanding.

Of course, such a pure understanding, which is usually identified with the divine or intuitive intellect, is a mere fiction for Kant, or as he puts it, a “problematic concept.” Nevertheless, this does not render reference to it otiose, since Kant's real point is that the human understanding proceeds *as if* it were such a pure understanding, whenever its thought outreaches the limits imposed by sensibility. And it can do this because understanding and sensibility make distinct contributions to cognition, each being governed by its own conditions. Thus, in spite of what some of Kant's formulations suggest, the thought (by the pure understanding) of things as they are in themselves is not *completely* empty; it has a certain content. At the same time, however, this content is of a merely logical nature, since it is derived from a use of the categories apart from the sensible conditions (schemata) that realize them; and such a use for Kant is itself merely logical (rather than real). Otherwise expressed, a consideration of things by means of pure categories (as some putative pure understanding might think them) is capable of yielding analytic judgments concerning the implications of the concepts of things so considered, but not synthetic *a priori* knowledge of the things themselves.<sup>45</sup>



Moreover, from this we can see a bit more concretely just what is wrong with Langton's proposed route to humility via receptivity, without any appeal to idealism. The basic problem is her neglect of the discursivity thesis, in terms of which the specific contribution of human sensibility is to be understood. First, in neglecting this thesis, she is inevitably led to misconstrue the epistemic role


of receptivity or affection for Kant. Instead, interpreting the latter in strictly causal (rather than epistemic) terms, she fails to see how for Kant (in contrast to the empiricists) receptivity brings along its own form or manner of being affected. Second, as a result of neglecting the same thesis, she mistakenly takes the *thought* of things as they are in themselves, by the pure (unschematized) concept of substance, to be at least minimally informative about the *nature* of such things. Thus, on her view we can know *that* things as they are in themselves are genuine substances consisting of merely intrinsic properties, with the humility consisting only in our inability to know what these properties are. But this is to ignore completely the purely analytic nature of such claims, which, as such, are capable of illuminating how a discursive intelligence is constrained to conceive of things so considered, but not their real nature (the latter requiring intuition as well as thought).




Conversely, focusing on the discursivity thesis makes it clear that true Kantian humility cannot bypass transcendental idealism, because it is a consequence of this thesis that the thought of things as they are in themselves abstracts from an essential condition of human cognition. Since thought can abstract from this condition, it can *consider* things as they are in themselves, that is, form a concept of things so considered. But for the very same reason, such thought does not amount to genuine (synthetic) knowledge, though it is not thereby necessarily trivial or tautologous.

The next two chapters will expand upon these claims, first approaching transcendental idealism obliquely by considering it in relation to the transcendental realism to which Kant opposes it, and then analyzing in some detail such central concepts as the thing in itself (or as it is in itself), the noumenon, the transcendental object, and affection. Before we turn to these topics, however, it may prove instructive to consider the objection that an “anodyne” interpretation of transcendental idealism amounts to a trivialization of it.

Although this criticism has been raised by both Guyer and Langton, it will suffice to consider the latter’s version, which is more directly related to the issues that have been discussed here. Her major complaint seems to be that, on my reading, Kantian humility is trivialized, because it is reduced to an analytic consequence of the definition of “a thing considered as it is in itself.” Since to consider a thing in this way is just to consider it in abstraction from the conditions of our cognition, it becomes trivially true that we can have no knowledge of things so considered. But she also thinks that this cannot be correct as an interpretation of Kant, since he viewed the humility thesis as a major philosophical discovery (not a trivial inference), and a “depressing” one at that.<sup>46</sup>

That Kant viewed the transcendental distinction between things as they appear and the same things as they are in themselves as a major philosophical discovery is undeniable. It is likewise undeniable that he regarded the limitation of cognition to the former as a consequence of this distinction. Nevertheless, it would be more accurate to say that he viewed this limitation as liberating or therapeutic rather than as depressing. As we shall see in connection with the discussion of the Transcendental Dialectic, Kant thought that this limitation provides the only means to avoid being deceived by an illusion that is inherent in the very nature of human reason (transcendental illusion). More to the present point, however, that a conclusion follows analytically, once a distinction is in place, does not render it trivial. This would follow only if the distinction in question were itself obvious or trivial. But this is far from the case with the transcendental distinction, which, it will be argued, rests upon a radical reconceptualization of human knowledge as based on *a priori* conditions (epistemic conditions). 

In sum, rather than being, in Guyer's dismissive phrase, "an anodyne recommendation of epistemological modesty," transcendental idealism, as here understood, is a bold, even revolutionary, theory of epistemic conditions. This, of course, is not to prejudge either the validity of this interpretation or the viability of transcendental idealism so understood. It is, however, to claim, pace both Guyer and Langton, that there is nothing trivial about it. 



## Chapter 2 Transcendental Realism and Transcendental Idealism

The previous chapter approached transcendental idealism more or less directly. The goal was to define this idealism, at least in a preliminary fashion, by locating its foundations in the specific conditions of discursive cognition. It was claimed that this location both provided a warrant for drawing the transcendental distinction between things considered as they appear and as they are thought in themselves, and justified the limitation of knowledge to the former (what Langton calls “Kantian humility”). The present chapter takes a more indirect route to the same end. The strategy is to interpret transcendental idealism by means of the transcendental realism that Kant opposes to it. This approach is based on the hermeneutical principle that often the best way to understand a philosophical position is to become clear about what it denies. It derives added justification from the fact that Kant appears to regard these two forms of transcendentalism as mutually exclusive and exhaustive metaphilosophical alternatives.<sup>1</sup> The chapter is divided into three parts: the first considers transcendental realism in its various guises; the second investigates the nature of tran-



scendental idealism, viewed as the single alternative to this realism; and the third considers two objections to this interpretation of transcendental idealism.

## I. THE NATURE OF TRANSCENDENTAL REALISM

The first difficulty confronting the strategy adopted here is that the significance attributed to transcendental realism seems to be belied by the relative paucity of references to it in the text. One would normally expect to find a conception of such alleged importance analyzed in great detail and subjected to a searching critique. But, apart from the bald claim that such a realism would undermine the possibility of both nature and freedom (A543 / B571), it is explicitly referred to in only two other places in the *Critique*. Both are in the Transcendental Dialectic, and in each case Kant contrasts it with transcendental idealism. The first is in the first-edition version of the Fourth Paralogism. Kant's concern there is to refute empirical idealism, which he contrasts with his own transcendental version. In this context he writes:

I understand by the **transcendental idealism** of all appearances the doctrine that they are all together to be regarded as mere representations and not as things in themselves, and accordingly that time and space are only sensible forms of our intuition, but not determinations given for themselves or conditions of objects as things in themselves. To this idealism is opposed **transcendental realism**, which regards space and time as something given in themselves (independent of our sensibility). The transcendental realist therefore interprets outer appearances (if their reality is conceded) as things in themselves, which would exist independently of us and of our sensibility and thus would also be outside us according to pure concepts of the understanding. It is really this transcendental realist who afterwards plays the empirical idealist; and after he has falsely presupposed about objects of the senses that if they are to exist they must have their existence in themselves even apart from sense, he finds that from this point of view all our representations of sense are insufficient to make their reality certain. [A369]

Kant is here arguing that transcendental realism leads to empirical idealism, which is the doctrine that the mind can have immediate access only to its own ideas or representations, that is, the familiar Cartesian-Lockean theory of ideas. His basic point is that, because this form of realism regards "outer appearances" (spatial objects) as things in themselves, it is forced to concede that the existence of such objects is problematic, since the mind has no immediate access to them. Transcendental realism is thus presented as the source of the pseudo-

problem of the external world and of the typically Cartesian version of skepticism associated with it.

The second passage is from the Antinomy of Pure Reason. There Kant defines transcendental idealism as the doctrine that “all objects of an experience possible for us are nothing but appearances, i.e., mere representations, which, as they are represented, as extended beings or series of alterations, have outside our thoughts no existence grounded in itself.” In contrast to this, the transcendental realist is said to make “these modifications of our sensibility into things subsisting in themselves, and hence makes mere representations into things in themselves” (A490–91 / B518–19).

Both of these passages indicate that the defining characteristic of transcendental realism is its confusion of appearances, or “mere representations,” with things in themselves. The first limits this charge to objects of “outer perception” (empirically external, spatial objects), although it does connect this realism with the conception of time as well as space as given in themselves, independently of our sensibility. This emphasis on space and outer experience no doubt reflects Kant’s concern at that point with empirical idealism and its connection with transcendental realism. The second passage, which does not reflect this particular concern, goes somewhat further by presenting transcendental realism as the view that considers *all appearances*, those of inner sense as well as those of outer sense, as if they were things in themselves. Clearly, the latter passage expresses Kant’s considered view on the subject. Since it is a central tenet of the *Critique* that inner as well as outer sense present us with objects as they appear, not as they are in themselves, transcendental realism manifests itself as much in a confused view of the former as of the latter.

This of itself should make it clear that the usual interpretation of transcendental realism as equivalent to the scientific realism of the Cartesians and Newtonians (roughly what Berkeley meant by “materialism”) is far too narrow.<sup>2</sup> Although Kant only infrequently makes use of the expression, he repeatedly accuses philosophers of a variety of stripes of treating appearances as if they were things in themselves or, equivalently, of granting “absolute” or “transcendental” reality to appearances.<sup>3</sup> Indeed, at one place in the *Critique* he terms this confusion the “common prejudice” (A740 / B768), while at another he refers to the “common but fallacious presupposition of the absolute reality of appearances” (A536 / B564). Moreover, this claim is found in even stronger form in other texts. In fact, he goes so far as to assert that prior to the *Critique* the confusion was unavoidable (Fort 20: 287; 377) and even that “until the crit-

ical philosophy all philosophies are not distinguished in their essentials" (Fort 20: 335; 413).

Such statements support the contention that the transcendental distinction between appearances and things in themselves or, more properly, between things as they appear and the same things as they are in themselves, functions as the great divide in the Kantian conception of philosophy. Only the "critical philosophy" has succeeded in getting this distinction right. As a result, despite their many interesting differences, all of the others are at bottom nothing more than variant expressions of the same underlying confusion.



Admittedly, such a sweeping claim, by which all previous and most succeeding philosophies are painted with one brush, seems highly suspicious on the face of it. Accordingly, before considering it in detail, it may be useful to keep in mind that Kant explicitly made a parallel claim regarding the significance of his contribution to the subject in the area of moral philosophy. Thus, in both the *Groundwork* and the *Critique of Practical Reason*, in introducing autonomy as the supreme principle of the possibility of the categorical imperative, Kant contrasts the principle of autonomy with that of heteronomy and maintains that all previous moral theories were committed to the latter.<sup>4</sup> The present suggestion, then, is that transcendental realism, understood as the point of view that systematically identifies appearances with things in themselves, be assigned the same role in Kant's theoretical philosophy that he assigned to heteronomy in his moral philosophy. In other words, it constitutes the common assumption, standpoint, prejudice, or confusion shared by all philosophers who do not adhere to the critical view.<sup>5</sup>



### A. Some Varieties of Transcendental Realism

The best way to test this suggestion is to see the extent to which it is applicable to various "noncritical" philosophies. It should be noted, however, that in so doing we shall explicitly be viewing these philosophies through Kantian spectacles. The question is not whether the charge that they confuse appearances with things in themselves is "fair" according to some independent standard of evaluation. It is rather whether, given Kant's assumptions, it is reasonable to view these philosophies in such a manner.

To begin with, we have already seen that Kant maintains that empirical idealism is a form of transcendental realism, which arises from the recognition of the fact that the human mind has no direct access to the putatively "real" things, that is, to physical objects construed as things in themselves. This recog-

nition, in turn, leads to the claim of Descartes and his followers that the only objects of which we are immediately aware are ideas in the mind. Such idealism, together with its skeptical consequences, is, therefore, the result of an implicit commitment to transcendental realism. Kant's first-edition version of the *Refutation of Idealism* turns on this point. As he succinctly puts the matter:

If we let outer objects count as things in themselves, then it is absolutely impossible to comprehend how we are to acquire cognition of their reality outside us [*ausser uns*], since we base this merely on the representation, which is in us [*in uns*]. For one cannot have sensation outside oneself, but only in oneself, and the whole of self-consciousness therefore provides nothing other than merely our own determinations. [A378]

At first glance this seems reminiscent of Berkeley's critique of "materialism," and it has frequently been taken in just this way.<sup>6</sup> On this reading, Kant, like Berkeley, succeeds in avoiding skepticism only by identifying the "real" with the immediate objects of consciousness. It should be apparent from our preliminary discussion in the previous chapter, however, that such a reading constitutes a gross distortion of Kant's position, since it ignores its explicitly transcendental thrust.

This thrust is most clearly evident in Kant's disambiguation of the key terms '*in uns*' and '*ausser uns*'. As he points out, these can be taken in either an empirical or a transcendental sense (A373). Taken in the former way, they mark a contrast between how objects are experienced: either as temporally located objects of inner sense or as extended, spatially located objects of outer sense. Taken in the latter way, they contrast two manners in which objects can be considered in relation to the conditions of human sensibility. From this transcendental standpoint, things may be viewed as *in uns* (or even as "mere representations") insofar as they are regarded as subject to the sensible conditions of cognition (space and time) or, equivalently, as phenomena or objects of possible experience. They are regarded as *ausser uns* insofar as they are thought independently of these conditions "as they are in themselves."<sup>7</sup>

Viewed in the light of this distinction, the form of transcendental realism that results in empirical or skeptical idealism is guilty of a kind of category mistake. Specifically, it takes the merely empirically external (spatial) objects to be *ausser uns* in the transcendental sense. Or, more properly, it fails to distinguish between these two senses of being *ausser uns*. And from this the transcendental realist concludes correctly that the human mind has no direct cognitive access to objects so considered. The mistake here is not in assuming that things exist

independently of their relation to the conditions of human sensibility (Kant likewise assumes this); it is rather in assuming that things existing in this way retain their spatiotemporal properties and relations. Thus, by linking Cartesian skeptical idealism to transcendental realism, Kant shows not only how transcendental idealism provides the solution, but also how it supplies the means for diagnosing the problem.

Nevertheless, not all forms of transcendental realism are committed to empirical idealism and the skepticism it engenders. A prime example of a transcendently realistic mode of thought that is not is that of the Newtonians or “mathematical students of nature.” As I have already suggested, their conception of absolute space and time amounts to treating the latter as ontological (rather than epistemic) conditions, which is equivalent to viewing them (as well as the things in them) as *ausser uns* in the transcendental sense.

We shall see that similar considerations apply also to Leibniz, whom Kant explicitly accuses of taking appearances as things in themselves (A264 / B320). In order to test the thesis that the label “transcendental realism” is applicable to all noncritical philosophies, however, the most pertinent examples are obviously the phenomenalist views of Berkeley and Hume.<sup>8</sup> For if even these thinkers can be shown to have confused appearances with things in themselves, it can be claimed with some justice that the confusion is virtually universal.

To begin with, Kant views Berkeley’s “dogmatic idealism” as in a certain sense the logical outcome of the absurdities inherent in the Newtonian conceptions of absolute space and time as ontological conditions.<sup>9</sup> As he puts it in a second-edition addendum to the Transcendental Aesthetic:

For if one regards space and time as properties that, as far as their possibility is concerned, must be encountered in things in themselves, and reflects on the absurdities in which one then becomes entangled, because two infinite things that are neither substances nor anything really inhering in substances must nevertheless be something existing, indeed the necessary condition of the existence of all things, which also remain even if all existing things are removed; then one cannot well blame the good Berkeley if he demotes bodies to mere illusion; indeed even our own existence, which would be made dependent in such a way on the self-subsisting reality of a non-entity such as time, would be transformed along with this into mere illusion; an absurdity of which no one has yet allowed himself to be guilty. [B70–71]<sup>10</sup>

Since we have seen that the Newtonian conception is itself transcendently realistic, it follows that Berkeley’s denial of material substance, which Kant dismissively glosses as “demot[ing] bodies to mere illusion,” should be viewed as at

least an indirect offshoot of such realism. As such, it stands to Newtonian absolute space and time roughly as empirical idealism stands to Cartesian *res extensa*. In other words, it is a form of subjectivism or idealism to which one is driven on the basis of certain transcendently realistic assumptions.

Further consideration, however, suggests that Berkeley's position is not merely an indirect offshoot of transcendental realism; it is also itself transcendently realistic, because, like other forms of such realism, it regards Kantian appearances as *ausser uns* in the transcendental sense. Admittedly, this may seem paradoxical in the extreme, since on Kant's scheme Berkeleyan ideas are *in uns* in the empirical sense. But the paradox disappears if one keeps in mind that to be *ausser uns* in the transcendental sense just means to exist independently of the conditions of human sensibility. Accordingly, there is no incompatibility between being *in uns* in the empirical and *ausser uns* in the transcendental sense. In fact, this is precisely the status that Kant assigns to inner appearances or objects of inner sense. The problem, though, is that Berkeley's idealism inverts the true order of things by attributing this status to *outer* appearances.

Although this analysis goes beyond what Kant says about Berkeley, it finds strong confirmation in a similar claim that he makes about Hume (which seems equally applicable to Berkeley). The crucial passage occurs in the *Critique of Practical Reason*, where, by way of summarizing some of the essential tenets of the first *Critique*, Kant reflects:

When Hume took objects of experience as things in themselves (as is done almost everywhere), he was quite correct in declaring the concept of cause to be deceptive and a false illusion; for, as to things in themselves and the determinations of them as such, it cannot be seen why, because something, A, is posited, something else, B, must necessarily be posited also, and thus he could certainly not admit such an *a priori* cognition of things in themselves. (KpV 5: 53; 182)

Since Kant was well aware that Hume characterized the objects of human awareness as "impressions," we are led to ask why he should claim that Hume viewed them as things in themselves. Kant's point, of course, is not that Hume thought he was doing anything of the sort but, rather, that this is what his position amounts to, when considered from a transcendental perspective. Moreover, for Kant this is the consequence of Hume's failure to recognize the existence of *a priori* forms of sensibility through which the mind receives its impressions.<sup>11</sup> Since, as the passage goes on to suggest, Hume did not recognize any such *a priori* forms, he could not acknowledge the possibility of any *a priori* rules of synthesis through which impressions are brought to the unity of

consciousness. In the absence of such rules, however, there is no reason why, given object (or impression) A, something else, object (or impression) B, must likewise be given; and this, as Kant sees it, is the source of Hume's skeptical doubts concerning causality.<sup>12</sup>

Although this raises important questions concerning Kant's critique of Hume, with which we shall be concerned later, our present focus must be limited to the implications of Kant's analysis for the understanding of transcendental realism. Moreover, here the implication is clear. Notwithstanding their subjectivist accounts of the objects of human awareness, both Berkeley and Hume may be said to view appearances *as if* they were things in themselves, because they deny any *a priori* contribution of sensibility to the cognition of these appearances. Accordingly, they regard spatiotemporal objects (Kantian appearances) as *ausser uns* in the transcendental sense, while at the same time treating them (erroneously from Kant's point of view) as *in uns* in the empirical sense. Thus, they are both transcendental realists.

## B. Transcendental Realism and the Theocentric Model of Knowledge

When we combine this result with that of the previous chapter, it seems clear that what all forms of transcendental realism have in common may be negatively expressed as a failure or, to put it less tendentiously, a refusal, to recognize that human cognition rests on *a priori* conditions of sensibility, which structure the way in which the mind receives its sensory data. Moreover, if the earlier analysis is correct, this is tantamount to a failure to acknowledge the discursive nature of human cognition. Thus, transcendental realism goes hand in hand with the rejection of the discursivity thesis.

This rejection is reflected in the downgrading of conceptual representation by both rationalism and empiricism. The underlying complaint is that, in virtue of its generality, such representation is at best partial and abstract; and, as such, it fails to grasp objects in their full concreteness.<sup>13</sup> From the rationalist side, this is expressed in the contrast drawn by Spinoza between the second and third kinds of cognition (discursive cognition or *ratio* and intuitive cognition or *scientia intuitiva*), of which only the latter is judged capable of grasping the essence of individual things.<sup>14</sup> Among the empiricists, it largely takes the form of a worry about abstract ideas, which, quite apart from the psychological question of the possibility of forming them, are likewise deemed inadequate to apprehend an object as it is in itself.<sup>15</sup>

Nevertheless, it will not suffice to define transcendental realism in purely



negative terms, since it fails to indicate the sense in which this realism constitutes a consistent standpoint, shared by a wide variety of distinct philosophical positions. Thus, for all that we have seen so far, it might seem that it is nothing more than a label invented by Kant to encompass everything to which he was opposed. But if this is the case, it further seems that the project of attempting to understand transcendental idealism in terms of its contrast to such realism is doomed to failure.

Consequently, it is crucial to characterize transcendental realism in a positive manner, and the suggestion is that this is best accomplished by defining it in terms of a commitment to a theocentric paradigm or model of cognition.<sup>16</sup> To reiterate a point made previously, since cognition (of whatever sort) requires that its object somehow be “given” to the mind, the denial of discursivity requires the assumption that the objects themselves (and as they are in themselves), not merely the data for thinking them, be so given. Thus, if, as empirical idealism avers, objects are not given (but only inferred), skepticism inevitably ensues. But since the only kind of intuition that could supply the objects themselves is intellectual, which is traditionally thought to characterize a divine or infinite intellect, it follows that transcendental realism is committed to a theocentric paradigm in virtue of its denial of discursivity. In fact, these are merely two sides of the same coin.

This claim, however, must be qualified in at least two essential respects. First, the point is not that Kant either thought that transcendental realism is committed to the existence of an intuitive intellect or that he assumed that all such realists assert the cognizability of things as they in themselves in the sense in which he understands the notion. It is rather that the idea of such an intellect functions as an implicit norm in the light of which human cognition is analyzed and measured. Since, *ex hypothesi*, such an intellect cognizes things as they are in themselves, it follows that any account of human cognition that appeals to this model (even if only implicitly) also assumes that its proper objects are things as they are in themselves. Second, Kant is not suggesting that the transcendental realist must hold that human beings actually possess intellectual intuition, or even some pale imitation thereof. Although there may be hints of such a view in certain rationalists, it is totally antithetical to empiricism in any form. The main point is rather that this realism considers our *sensible* intuition *as if it were intellectual*, because it tacitly assumes that, insofar as our intuition acquaints us with objects at all, it acquaints us with them as they are in themselves.

The theocentric model, with its ideal of an eternalistic, God’s-eye view of

things, is the common heritage of the Platonic tradition, but it is particularly evident in the great rationalists of the seventeenth century.<sup>17</sup> One thinks in this connection of Malebranche, who claimed that we “see all things in God,” and again of Spinoza, who maintained that the goal of human cognition is to view things *sub specie aeternitatis*.<sup>18</sup> It is also central to Leibniz, however, and, as I shall argue, provides the key to understanding both his form of transcendental realism and Kant’s critique thereof.

Moreover, in spite of their essentially psychological orientation, it is clear from their views on conceptual representation that the empiricists were also committed to this model. Although most apparent in Berkeley, who was something of a Platonist, it is also equally true of Locke and Hume. But since the transcendently realistic dimension of Hume’s thought has already been noted, the discussion will focus on Locke, in whom the connection between this model and his views on conceptuality is particularly perspicuous. Finally, in an effort to underscore the prevalence of this model and to provide a further basis for understanding the nature of Kant’s “Copernican revolution,” it will be shown that it also underlies Kant’s own pre-critical thought.

*Leibniz.* Leibniz’s appeal to the theocentric model is quite explicit and has often been noted in the literature.<sup>19</sup> Following Augustine and Malebranche, Leibniz depicts the divine understanding as the realm of eternal truths, and he claims that it is there that one finds “the pattern of the ideas and truths which are engraved in our souls.”<sup>20</sup> This is not to say that the human mind for Leibniz is infinite, or that it is somehow capable of thinking “God’s thoughts.” On the contrary, he constantly emphasizes the insurmountable limits of human knowledge and explains these in terms of the confusedness of our representations, which is itself seen as a consequence of our finitude. The point, however, is not that human knowledge is infinite, or even often adequate, for Leibniz; it is rather that it approaches adequacy as it approaches divine knowledge. Thus, despite the infinite difference in degree or scope, Leibnizian rationalism assumes a commensurability or similarity in kind between human and divine knowledge.<sup>21</sup>

This assumption underlies Leibniz’s claim that in any true proposition the predicate is contained in the concept of the subject. Leibniz’s adherence to this principle leads him to regard demonstration as requiring reduction to identity. He thinks that this is quite possible for arithmetical propositions and possible, at least in principle, for the axioms of Euclidean geometry. Moreover, he holds that this principle is applicable not only to necessary truths or “truths of reason,” which are true in all possible worlds, but also to contingent truths or



“truths of fact,” which hold only in the actual world. As Leibniz puts it at one point, this is because “it is the nature of an individual substance or complete being to have a concept so complete that it is sufficient to make us understand and deduce from it all the predicates of the subject to which the concept is attached.”<sup>22</sup> Since the complete concept of an individual substance involves an infinity of elements, and since a finite mind is incapable of infinite analysis, the human intellect can never arrive at such a conception. As a result, it cannot demonstrate or deduce truths of fact. Nevertheless, such truths remain cognizable in principle, that is, for God, who is capable of an intuitive grasp of the infinite. Expressed in Kantian terms, this means that all propositions are ultimately analytic and that the syntheticity of truths of fact is merely a function of the limits of analysis, not of the nature of the propositions themselves.



These considerations better enable us to grasp the main outlines of Kant's critique of Leibniz and to understand the claim that the latter took appearances for things in themselves. Much of Kant's quarrel with Leibniz and his followers turns on the closely related conceptions of sensibility and appearance. By and large, Kant defines his philosophy vis à vis Leibniz's in terms of their different understanding of these conceptions. He claims that Leibniz and his followers “falsified” both conceptions, and he sees this as the direct result of their understanding of the distinction between the “sensible” and the “intelligible.” Instead of viewing the difference between these two elements of human cognition as “transcendental,” that is, as a difference of origin, content, and kind, they regard it as merely “logical,” that is, as a difference of degree of clarity and distinctness of the representations.<sup>23</sup> All of this is captured by the claim that Leibniz (here contrasted with Locke), “intellectualized appearances” (A271 / B327). To “intellectualize appearances” for Kant is to abstract from their irreducibly sensible (spatiotemporal) character. But since this character is a defining feature of a Kantian appearance, while independence of it is a defining feature of a thing as it is in itself, it can easily be seen that this is equivalent to mistaking the former for the latter.

In his response to Eberhard, Kant makes it clear that the heart of the difficulty with Leibnizianism is that it fails to recognize that human sensibility has its own *a priori* forms or conditions (space and time), which serve to determine positively the nature and relations of the objects of human experience.<sup>24</sup> That is why Leibnizians regard sensible (perceptual) knowledge of appearances merely as a confused version of the purely intellectual knowledge obtained by God. Consequently, all of the sensible components of human experience, including spatiotemporal relations, are deemed reducible (for God) to the purely intellectual

(logical) determinations that pertain to things in themselves (monads). This view of sensible cognition is, however, the logical consequence of Leibniz's appeal to the theocentric model of knowledge and thus of his transcendental realism. It is, therefore, the latter that is the real object of Kant's critique.<sup>25</sup>

*Locke.* Although not as prominent, Locke's appeal to the theocentric model is just as real as Leibniz's. Perhaps the best example of this is his much discussed distinction between nominal and real essence. By the nominal essence of a substance, really of a "sort," Locke understands the complex idea of that sort. This idea, like all general ideas for Locke, is due to the "workmanship of the understanding," which forms it on the basis of the experience of a number of resembling particulars. Such ideas therefore constitute the senses of sortal terms. The real essence, by contrast, is the inner nature or "real constitution" of a thing. Locke uses the example of gold to illustrate this distinction. "The nominal essence of gold," he tells us, "is that complex idea the word gold stands for, let it be, for instance, a body yellow of a certain weight, malleable, fusible and fixed;" whereas its real essence is characterized as "the constitution of the insensible parts of that body, on which those qualities and all other properties of gold depend."<sup>26</sup>

As products of the human understanding, sortal concepts or nominal essences are clear exemplars of conceptual representations. But what makes this particularly interesting for our purposes is that Locke correlates the distinction between the two kinds of essence with the distinction between divine and human knowledge. A nice illustration of this is his analysis of the "essence" of man. After briefly categorizing those features that are contained in the complex ideas constituting the nominal essence of man, Locke writes in a memorable passage:

The foundation of all those qualities which are the ingredients of our complex idea, is something quite different: and had we such a knowledge of that constitution of man, from which his faculties of moving, sensation, and reasoning, and other powers flow, and on which his so regular shape depends, as it is possible angels have, and it is certain his Maker has, we should have a quite other idea of his essence than what now is contained in our definition of that species, be it what it will: and our idea of any individual man would be as far different from what it is now, as is his who knows all the springs and wheels and other contrivances within the famous clock at Strasburg, from that which a gazing countryman has of it, who barely sees the motion of the hand, and hears the clock strike, and observes only some of the outward appearances.<sup>27</sup>

Knowledge of real essence is here explicitly equated with the knowledge that our "Maker has." Human cognition, by contrast, is limited to "some of the outward appearances of things." Clearly, then, the latter is judged by the ideal standard of divine knowledge and found wanting. Moreover, what makes it inferior is precisely its conceptual nature, which limits its scope to manifest resemblances, the surface, rather than the deep structure, of things. Locke's agnosticism is mitigated, however, by his characteristic insistence that the nominal essences produced by the understanding and the classifications based upon them are sufficient for our needs. As he eloquently expresses it in the Introduction to the *Essay*, "The candle that is set up in us shines bright enough for all our purposes."<sup>28</sup> These purposes include not only knowledge of God and of our duty, but also what Locke calls "the conveniences of life."<sup>29</sup> His point is that our classification of things into sorts and, more generally, our empirical cognition, suffices to attain these "conveniences," even though it does not acquaint us with the true nature of things. Locke, therefore, combines his appeal to the theocentric model with an essentially pragmatic account of empirically based conceptual cognition. In this respect his position is not far from that of rationalists like Descartes and Malebranche.

The primary difference between Locke and the rationalists on this score is that Locke tends to conceive of fully adequate or divine knowledge as basically more of the same; that is to say, he regards it as if it were perceptual in nature, albeit an idealized perception, liberated from any need to rely on general ideas because of greatly expanded powers, for example, "microscopical eyes."<sup>30</sup> This is presumably what Kant had in mind, when, in contrasting Locke with Leibniz, he remarks that Locke "sensitized the concepts of understanding," and that he viewed sensibility "as immediately related to things in themselves" (A271 / B327).

As the connection between sensibility and things in themselves indicates, Locke's "sensitization" of the concepts of the understanding is not to be viewed as indicating an abandonment of the theocentric model with its cognitive ideal of intellectual intuition. On the contrary, what Kant regards as an intellectual intuition, that is, a direct and complete acquaintance with an object as it is in itself (unmediated by any conceptual representation) is construed by Locke as perceptual in nature.<sup>31</sup> Thus, even though they interpret it in radically different ways, both Locke and Leibniz assume that human cognition is to be analyzed in light of the theocentric model.<sup>32</sup>

*The Pre-critical Kant.* Perhaps the most instructive example of an appeal to the theocentric model is provided by Kant himself. Indications of this appeal can

be discerned in virtually all of his pre-critical writings, but for illustrative purposes we can limit our consideration to “A New Exposition of the First Principles of Metaphysical Cognition” (1755), a work that reflects a stage in his development at which Kant philosophized very much in a Leibnizian mold. Thus, in support of the claim that the principle of identity is the first principle of all truths, the young Kant writes:

[S]ince all our reasoning amounts to uncovering the identity between the predicate and the subject, either in itself or in relation to other things, as is apparent from the ultimate rule of truths, it can be seen that God has no need of reasoning, for, since all things are exposed in the clearest possible way to his gaze, it is the same act of representation which presents to his understanding the things which are in agreement and those which are not. Nor does God need the analysis which is made necessary for us by the night which darkens our intelligence. [ND 1: 391; 10]

Kant here expresses as clearly as one might wish his commitment to the theocentric model with its ideal of a non-conceptual, purely intuitive cognition. Being finite cognizers, we are forced to have recourse to analysis (and, therefore, conceptualization) in order to grasp the identities that the divine intellect recognizes immediately. Moreover, Kant’s commitment to this model is revealed not only in this formulation of the ideal of cognition but also in some of the central arguments of the work. Two examples should suffice to make this clear. The first occurs within Kant’s argument for the existence of God as the ground of the possibility and hence of the essence of things. In developing this argument, Kant appeals to the example of the essence of a triangle:

For the essence of a triangle, which consists in the joining together of three sides, is not in itself necessary. For what person of sound understanding would wish to maintain that it is in itself necessary that three sides should always be conceived as joined together? I admit, however, that this is necessary for a triangle. That is to say: if you think of a triangle, then you must necessarily think of three sides. And that is the same as saying: “If something is, it is”. But how it comes about that the concepts of sides, of a space to be enclosed, and so forth, should be available for use by thought; how, in other words, it comes about that there is, in general, something which can be thought, from which there then arises, by means of combination, limitation and determination, any concept you please of a thinkable thing—how that should come about is something which cannot be conceived at all, unless it is the case that whatever is real in the concept exists in God, the source of all reality. [ND 1: 395–96; 16–17]

The second example occurs in connection with the claim that the principle of the coexistence of substances is to be located in the divine intellect. In support of this contention, Kant reflects:

[I]t has to be admitted that this relation depends on a community of cause, namely on God, the universal principle of beings. But it does not follow from the fact that God simply established the existence of things that there is also a reciprocal relation between these things, unless the self-same scheme of the divine understanding, which gives existence, also established the relations of things to each other, by conceiving their existences as correlated with each other. It is most clearly apparent from this that the universal interaction of all things is to be ascribed to the concept alone of this divine idea. [ND 1: 413; 41]

The primary import of these passages lies in the light they shed on the elements of continuity and change in Kant's thought. Both the "pre-critical" and the "critical" Kant were concerned with the determination of the conditions of possibility, though these conditions are understood in quite different ways. In the first passage, the question at issue is the nature of the ground or the condition of the possibility of three straight lines enclosing a space. The answer of the young Kant is that it is grounded in its conceivability by the divine intellect. By contrast, in his account of mathematical possibility in the *Critique*, Kant argues that the impossibility of two straight lines enclosing a space is based upon the conditions of the constructability of figures in space (A221 / B268), with these conditions being themselves determined by the nature of human sensibility.

The second passage is even more striking, for Kant poses the very same problem that he later deals with in the Analogies, namely, the ground of the unity of experience. In the *Critique* this unity is explained in terms of certain principles (the Analogies), which, as we shall see, function as the conditions of the possibility of the experience of a unified time order and express the necessary conformity of appearances to the schemata of the pure concepts of the understanding. Here, by contrast, the objects (substances) are held to conform necessarily to the schema of the divine intellect. The appeal to the divine intellect in this early essay thus fulfills much the same function as does the appeal to the human intellect in the *Critique*, which further suggests that what is generally characterized as Kant's "transcendental turn" may be plausibly regarded as a shift from a theocentric to an anthropocentric paradigm.<sup>33</sup>



## II. THE TRANSCENDENTAL NATURE OF KANT'S IDEALISM

It was argued in the preceding section that all noncritical philosophies, including that of the young Kant, can be regarded as transcendently realistic and that, as such, they share a commitment to the theocentric paradigm, which

goes together with a degradation of the discursivity that is the hallmark of a finite intellect. Since this encompasses philosophies of widely different kinds, transcendental realism cannot itself be defined in straightforwardly ontological, or even epistemological, terms. Instead, it was suggested that it must be understood in broadly metaphilosophical or metaepistemological terms as a “standpoint” or normative model with reference to which human cognition is analyzed and evaluated.

The remainder of this section will explore the implications of this result for the interpretation of transcendental idealism. The most important of these is that, like its counterpart, transcendental idealism must also be characterized as a metaphilosophical “standpoint,” rather than, as is usually done, as a metaphysical doctrine about the nature or ontological status of the objects of human cognition. Since the basic import of the Kantian position is most clearly reflected in Kant’s characterization of transcendental idealism as “formal” or “critical” and in the comparison of his procedure with that of Copernicus, we shall begin with a brief consideration of these. This should then put us in position to specify the fundamental difference between transcendental idealism and phenomenalism or an idealism of the Berkeleian sort.



### A. Transcendental Idealism as Formal

#### Idealism and the So-Called “Copernican Revolution”: Two Attempts at Clarification


In response to the pervasive misunderstanding and criticism of his idealism as it was formulated in the first edition of the *Critique*, Kant notes in the appendix to the *Prolegomena* that he now wishes transcendental idealism to be termed “‘formal’ or, better still, ‘critical’ idealism.” In so doing he hoped to distinguish it from both “the dogmatic idealism of Berkeley and the skeptical idealism of Descartes” (Pro 4: 375; 162–63). Moreover, in a note added in the second edition of the *Critique* to the previously cited definition of transcendental idealism, he remarks, “I have also occasionally called it formal idealism, to distinguish it from material idealism, i.e., the common idealism that itself doubts or denies the existence of external things” (B519).



Given the continued prominence of readings that interpret transcendental idealism as a version of the “common idealism,” Kant would have been well advised to follow more consistently his own terminological recommendation. This idealism is “formal” in the sense that it is a theory about the nature and the scope of the conditions under which objects can be cognized by the human mind.<sup>34</sup> It is “critical” because it is grounded in a reflection on the conditions



and limits of discursive cognition, not on the contents of consciousness or the nature of *an sich* reality. In both respects it differs radically from idealisms of the “common” sort, which are themselves forms of transcendental realism.

As I noted in the first chapter, the major source of the interpretive problem lies in Kant’s tendency to refer to the objects of human experience not only as “appearances” but also as “mere representations.” Nevertheless, even here careful attention to the text makes it possible to avoid the usual misunderstanding. Consider, for example, the characterization of transcendental idealism to which Kant appended the above-mentioned note. As we have seen, Kant there describes this idealism as the doctrine that “all objects of an experience possible for us, are nothing but appearances, i.e., mere representations, which, *as they are represented* [my emphasis], as extended beings or series of alterations, have outside our thoughts no existence grounded in itself” (A490–91 / B518–19). The apparent equation of appearances with “mere representations” in the main clause certainly suggests the common reading. The temptation to read it in this way disappears, however, once it is recognized that the italicized phrase in the subordinate clause refers back to the *objects* represented rather than to “appearances.” The claim, therefore, is not that these objects have no mind-independent existence (as one might maintain with regard to Berkeleian ideas); it is rather that such existence cannot be attributed to them in *the way in which they are represented*, that is, as spatiotemporal entities.<sup>35</sup> In short, such objects are *in uns* in the transcendental but not the empirical sense. Kant’s idealism is formal (rather than material) precisely because it allows for this distinction. 

Kant’s statement of what has come to be known as his “Copernican revolution” may be viewed as a second and closely related way in which he endeavored to clarify his idealism. This occurs in a famous passage from the Preface to the second edition of the *Critique*, in which Kant compares the “change in the way of thinking” [*Umänderung der Denkart*] that he has introduced into philosophy with the revolution in astronomy initiated by Copernicus (Bxvi). There is a considerable literature regarding the precise point of the comparison and the appropriateness of the Copernican analogy, the main point at issue being whether Kant has committed what is called the “anthropocentric fallacy” in his reading of Copernicus.<sup>36</sup> Fortunately, we need not concern ourselves with that issue here. The central question for us is rather how Kant’s own philosophical “revolution” is to be understood, which remains a question even if, as is frequently maintained, the analogy with Copernicus is not particularly apt. Kant describes his revolution thus:

Up to now it has been assumed that all our cognition must conform to the objects; but all attempts to find out something about them *a priori* through concepts that would extend our cognition have, on this presupposition, come to nothing. Hence let us once try whether we do not get farther with the problems of metaphysics by assuming that the objects must conform to our cognition. [Bxvi]

In view of the preceding analysis, it should be clear that Kant is here contrasting the “standpoints” of transcendental realism and transcendental idealism. To begin with, the assumption that “all our cognition must conform to the objects” is readily identifiable as the “common assumption” associated with transcendental realism. In consequence, the “objects” to which our cognition must conform are characterized as things in themselves. From this point of view, then, we can be said to know objects just to the extent to which our thought conforms to their “real” nature or, equivalently, to God’s thought of these same objects. On this model, Kant tells us, we cannot account for the possibility of [synthetic] *a priori* knowledge, because we cannot explain how the mind could “anticipate” any of the properties of objects so defined, which is required for *a priori* knowledge.<sup>37</sup> The problem is that this model assumes that all cognition rests ultimately upon a direct acquaintance with its object as it is in itself.

Although this is just what one would expect, given the normative idea of an intellectual intuition, in the case of finite, human cognition it entails that all knowledge must be *a posteriori*. In the *Prolegomena* Kant goes beyond this, however, suggesting that if the objects of human cognition were things as they are in themselves [*so wie sie an sich selbst sind*], it would not even be possible to account for *a posteriori* knowledge (Pro 4; 282; 78). Clearly, the latter represents Kant’s considered opinion, since his position is that transcendental realism, with its theocentric model, is incapable of explaining discursive cognition of *any* sort, not simply the *a priori* variety. That is why a philosophical revolution is necessary.

The contrary “Copernican” supposition that “objects must conform to our cognition” (*die Gegenstände müssen sich nach unseren Erkenntnis richten*), expresses the central tenet of transcendental idealism. In the previous chapter, this was taken to mean that objects must conform to the *conditions* under which we can alone represent them to ourselves as objects. This suggests both the notion of epistemic conditions, which was introduced as an expository device, and an anthropocentric model of cognition. Our present concern, however, is largely with the latter, which has been alluded to but not yet discussed.

Here everything depends on understanding the idea of such a model in a

normative sense. Otherwise Kant's position becomes essentially indistinguishable from that of the classical British empiricists, who, as the very titles of their major works indicate, likewise made a self-consciously anthropological turn. This is not to deny that Locke, Berkeley, and Hume were engaged, at least in part, in a normative enterprise. Clearly they were. As we have seen, however, in spite of their focus on the human understanding, human cognition, and human nature, these thinkers all analyzed cognition in the light of a theocentric norm. Thus, their central epistemological concern was to determine how human cognition stands with respect to such a norm, and in this they share common ground with the rationalists.

In sharp contrast to the procedure of the empiricists, to take the anthropological model in a normative sense is just to consider the human mind as the source of the rules or conditions through which and under which it can alone represent to itself an objective world. In Kant's terms, it is to say that the human understanding (suitably conditioned by sensibility) provides the "legislation [*Gesetzgebung*] for nature" (A126). Since our understanding is discursive (not intuitive), this entails that discursive cognition is elevated to the norm rather than degraded to a second-class form of cognition, as it inevitably is under the theocentric model.

## **B. Transcendental Idealism and Phenomenalism**

In light of the above, we are in a position to return to the question of the contrast between transcendental idealism and phenomenalism in general and Berkeleian idealism in particular. Jonathan Bennett's characterization of the nature of phenomenalism and its distinction from idealism provides a convenient starting point for this discussion. According to Bennett, phenomenalism is a theory about object language statements. It holds that all such statements are translatable into complex statements about sense data (including counterfactual hypotheticals). He further suggests that this is equivalent to the claim that "objects are logical constructs out of sense data." Idealism, by contrast, is characterized as the metaphysical view that "objects are collections of sense data." Bennett attributes the latter view to Berkeley.<sup>38</sup>

The first and most basic point to be made here is that phenomenalism, as Bennett describes it, is transcendently realistic in the same sense and for the same reasons as Berkeleian idealism. In spite of its conception of objects as "logical constructs," it treats the sensible data out of which "objects" are supposedly constructed as things in themselves. As a result, it is no more suitable for expli-

cating transcendental idealism than is Berkeleian idealism. In short, transcendental idealism is neither a theory about the translatability of object language statements into some more precise or primitive sense-datum language nor a theory about the ontological type (material object or collection of sense data) of the objects of human experience. As has been emphasized repeatedly, it is rather a theory about the *a priori* conditions and bounds of discursive cognition.

The issue can be clarified further by means of a comparison of Berkeley's analysis of statements about unperceived objects in *The Principles of Human Knowledge* with Kant's treatment of the same topic in the *Critique*. Berkeley offers two distinct analyses of propositions of the form: *x* exists, although *x* is not currently being perceived by myself or by another "created spirit." On one of these, *x* can be said to exist, if it is being perceived by God.<sup>39</sup> On the other, which is much closer to contemporary phenomenism, *x* can be said to exist if statements about *x* can be translated into hypotheticals of the form: if one were in position or had the proper instruments, and so forth, one would perceive *x*.<sup>40</sup> Both of these analyses are based upon the correlation between existence and perception, which is the hallmark of Berkeley's philosophy.

Kant's account of propositions about unperceived entities and events bears a superficial resemblance to Berkeley's second version, and therefore to phenomenistic accounts. Thus, he allows that we can perfectly well speak of inhabitants on the moon, even though no one has even seen them. But he goes on to note,

[T]his means only that in the possible progress of experience we could encounter them; for everything is actual [*wirklich*] that stands in one context with a perception in accordance with the laws of empirical progression. Thus they are real [*wirklich*] when they stand in an empirical connection with my real [*wirklich*] consciousness, although they are not therefore real [*wirklich*] in themselves, i.e., outside this progress of experience. [A493 / B521]

Moreover, Kant continues,

To call an appearance a real [*wirkliches*] thing prior to perception, means either that in the continuation of experience we must encounter such a perception, or it has no meaning at all. For that it should exist in itself without relation to our senses and possible experience, could of course be said if we were talking about a thing in itself. But what we are talking about is merely an appearance in space and time, neither of which is a determination of things in themselves, but only of our sensibility; hence what is in them (appearances) are not something in itself, but mere representations, which if they are not given to us (in perception) are encountered nowhere at all. [A493–94 / B522–23]



We can see from this that Kant, like both Berkeley and contemporary phenomenalism, translates first-order statements about unperceived entities or events into second-order statements about the possible perception thereof. But this superficial resemblance really masks the distinctive feature of the Kantian analysis, namely, the role given to *a priori* laws or principles. The “laws of the empirical progression,” or, as he calls them elsewhere, the “laws of the unity of experience” (A494 / B522), are nothing other than the Analogies of Experience. Without now entering into a discussion of these Analogies (this is the topic of chapter 9), the basic point is that, on a transcendently idealistic analysis, the claim that a certain entity or event is to be met with in the “progression of experience” is an elliptical way of affirming some lawful connection or “causal route” between the entity or event in question and present experience. It does not involve the postulation of a hypothetical mental episode in the history of some consciousness (whether human or divine).

The epistemic or transcendental thrust of Kant’s theory is brought out particularly clearly in the analysis of actuality [*Wirklichkeit*] in the Postulates of Empirical Thought. Kant there defines the actual as “that which is connected with the material conditions of experience (of sensation)” (A218 / B266). Because of the explicit reference to sensation, this definition seems to invite a phenomenalistic or even an idealistic reading (in the Berkeleian sense). Kant’s discussion of the postulate, however, suggests a different story. The claim that something is actual, we are told,

requires perception, thus sensation of which one is conscious—not immediate perception of the object itself the existence of which is to be cognized, but still its connection with some actual perception in accordance with the analogies of experience, which exhibit all real connection in an experience in general. (A225 / B272)

At first glance, this might suggest phenomenalism as Bennett defines it. To be sure, it rules out the extreme idealistic requirement that for an empirical object to be actual (real) it must be perceived, but it does seem to require the supposition that the object *could* be perceived, which is just the thesis of phenomenalism (with its appeal to counterfactuals). Nevertheless, this is not quite Kant’s position. Although he does hold that whatever is actual must be an object of possible perception, this is merely a consequence, not a criterion, of actuality. As the passage above indicates, the relevant criteria are provided by the Analogies of Experience, that is, by a set of *a priori* principles. The full critical position is that whatever can be connected with some given perception in accordance with these principles, or “laws of the empirical connection of appear-

ances,” is to be deemed “actual.” The appeal to perception or sensation here functions merely as the point of departure, which gives empirical content to the claim of actuality. The claim itself is not about any “subjective experiences.”

Kant’s illustration of this principle is also highly instructive. It concerns the hypothetical case of the perception of some magnetically attracted iron filings. Such a perception, he notes, would clearly justify the inference to the existence of some material responsible for this attraction. Moreover, it would do so even though our sensory apparatus is not adequate for the perception of this material. Admittedly, he then suggests that if our sense organs were more powerful or more refined we might be able to perceive it, which once again calls to mind phenomenalism’s appeal to counterfactuals (as well as Locke’s oblique reference to “microscopical eyes”). Kant, however, appeals neither to counterfactuals nor to the idea of a vastly improved sensory capacity. Instead, he remarks that “the crudeness [of our senses] . . . does not affect the form of possible experience in general. . . . Thus wherever perception and whatever is appended to it in accordance with empirical laws reaches, there too reaches our cognition of the existence of things” (A226 / B273). In other words, the meaningfulness of the reference to this magnetic material is not a function of the possibility of sufficiently improving our sensory apparatus, so as to enable us to have experiences that we are not at present able to have. It is rather a function of the connectibility of this material with our present experience in accordance with empirical laws and, ultimately, *a priori* principles.

The same point can be made with respect to the notion of a possible perception. As is already implicit in his *esse est percipi* principle, and as is perfectly manifest in his account of the *minimum sensible*, Berkeley’s account of possible perception is essentially psychological in nature. To be possible means to be actually perceivable. Accordingly, anything too small to be perceived, or below the *minimum sensible*, can simply be dismissed as impossible.<sup>41</sup> In sharp contrast to this, Kant defines the possibility of perception in terms of the conformity to rules, that is, to *a priori* principles. Thus, he writes:



[W]hat is required is only the progress from appearances to appearances, even if they should not yield any actual perception (if this perception is too weak in degree to become an experience for our consciousness), because despite this they would still belong to possible experience. (A522 / B550)

This passage almost seems as if it were written with Berkeley (or Hume) in mind. In any event, it nicely illustrates the radical difference between Kant’s transcendental or formal idealism and a phenomenalism or material idealism

of the Berkeleyian mold. The transcendental concept of appearance is linked here specifically to the notion of a possible experience. The latter notion, however, is defined in terms of conformity to a set of *a priori* conditions rather than in terms of the possibility of a perceptual state. Once again, then, we see that the appeal to such conditions, which are the conditions of discursive cognition, is the defining characteristic of transcendental idealism and that such idealism therefore has little in common with phenomenalism.<sup>42</sup>

### III. A REPLY TO TWO OBJECTIONS

In addition to the trivialization charge mounted by Guyer, Langton, and others, the present interpretation of transcendental idealism has been subject to a number of criticisms.<sup>43</sup> In concluding this discussion, we shall explore two of the most important of these: one of a substantive philosophical nature, the other largely a matter of interpretation. As representative of these two lines of criticism, we shall consider those of Jay Van Cleve and Karl Ameriks, respectively.

#### A. Van Cleve: One World or Two?

Van Cleve's objection is directed at the understanding of the transcendental distinction as holding between two ways of considering the same thing rather than between two ontologically distinct things. Suggesting that the texts are inconclusive on the issue (a matter to be taken up in the next chapter), he claims that the former alternative is untenable. The basic problem it confronts is explaining how the same thing could be both spatial and non-spatial or, more precisely, how it could be "considered as such."<sup>44</sup> Behind this way of formulating the problem is the recognition that such interpretations as the one offered here, which focus on the modifying phrase, do so in order to avoid the obvious contradiction in claiming that the same thing might *be* both spatial and non-spatial. Thus, he contends that we owe, but fail to provide, a general account of how modifiers might be thought to remove such a contradiction.

Following David Lewis, Van Cleve suggests three possible models for understanding this: "Square on the third floor, round on the fourth"; "Honest according to the *News*, crooked according to the *Times*"; "Tall compared to Ed, short compared to Fred."<sup>45</sup> As he correctly notes, the first two are obviously inadequate to model the Kantian distinction as here understood, since the first effectively transforms it into a distinction between two things, while the second makes one of the ways of considering things erroneous. Thus, we are left with



the third model as the “best bet.”<sup>46</sup> Central to this model is the distinction between relative and intrinsic properties, which means that, applied to Kant, we are to conceive of spatial properties, such as shape, as relative rather than, as they usually taken to be, intrinsic. In other words, shapes and such are disguised relations. And ignoring the fact that such a relational view is found in Leibniz, Van Cleve summarily dismisses it as untenable.

Nevertheless, in order to do justice to the position he is criticizing, Van Cleve stops to consider a model that was suggested in the original version of this book. This model, which was intended as an empirical illustration of a transcendental claim, involves the Newtonian conception of weight. According to this conception, bodies may be said to have weight only insofar as they stand in a relation of attraction and repulsion to other bodies. Hence, only insofar as a given body is “considered” in such a relation is a description including a reference to weight applicable to it. The intelligibility of this claim is in no way affected by the fact that bodies are always found to be in a relation of interaction with other bodies, so that “body as such” can never be an object of experience. The point is simply that bodies can very well be *conceived of*, though not *experienced*, apart from their relation to other bodies (Newton’s First Law of Motion is precisely about bodies so conceived of). Making allowance for the shift from the empirical to the transcendental level, it was suggested that much the same can be said about the distinction between things as they appear and the same things as they are in themselves. In this case also what we have is the distinction between a thing considered in a certain relation, in virtue of which it falls under a certain description, and the same thing considered in abstraction from this relation, and therefore not falling under this description.

Although Van Cleve is skeptical about this suggested model on the grounds that it still requires us to regard properties that are normally thought of as monadic as really relational, he does not dismiss it outright. Moreover, this is a good thing, since, as I have already noted, it corresponds to the Leibnizian view. Instead, Van Cleve focuses on the *disanalogy* between the relational understandings of weight and shape. His point is that whereas in the case of weight we can clearly understand the relation in question roughly as “being-pulled-to-a-certain-extent-by,” there is no comparable relation available for understanding shape.<sup>47</sup>

Van Cleve considers the most plausible candidate for such a relation to be that of “*appearing to us to have such-and-such a shape*.”<sup>48</sup> As he proceeds to argue, however, this commits Kant to the illusionist view that objects only *seem* to us to have spatial properties, though in reality they do not, which reduces to





the already discredited second model.<sup>49</sup> Thus, we are led by default to the “qualified two-world” view Van Cleve prefers.<sup>50</sup>

The response to this line of objection is twofold. First, Van Cleve is quite correct to see a problem here, since the relation in question is epistemic and, as such, differs significantly from the physical relation between body and weight. Nevertheless, his characterization of this relation is tendentious and conceals an important ambiguity. “*Appearing to us to have such-and-such a shape*” may mean either merely *seeming* to us to have it, much as the stick seen in the water seems to us to be bent, or as justifiably claimed to have it, qua considered in relation to the conditions under which it appears to beings with our forms of sensibility.<sup>51</sup>

Van Cleve’s entire critique rests on the assumption that it must be taken in the first sense. The only options he recognizes are:  $x$  really has property  $y$  (may be judged from a God’s-eye view to have it) or  $x$  only seems to us to have it. Clearly, this is the natural way to take the matter, since it appeals to our ordinary use of such language. Nevertheless, it also reflects a transcendently realistic position, which is likewise “natural” but completely bypasses the *transcendental* concept of appearance. Moreover, this is evidenced by the fact that Van Cleve understands the epistemic relation as fundamentally empirical in nature. Thus he refers to an object as having “such-and-such-a-shape” (e.g., round as opposed to square), which is an empirical matter, rather than as having shape (size or spatial location) at all, which is not. As a result, he begs the question concerning transcendental idealism as here interpreted.

Since Van Cleve might well reply at this point that this alternative transcendental conception of appearance, which is not to be identified with a mere seeming, remains opaque, it will be useful to provide another example, the consideration of which constitutes the second part of our reply. Although this concerns time rather than space, it provides a clearer illustration of the nature and force of Kant’s ideality thesis.

As is well known, traditional philosophical theologians generally hold that, as omniscient, God must have complete foreknowledge. Rather than, like finite beings, having to await events, God grasps in a timeless manner (through an “intellectual intuition”) everything that will ever happen. Usually, this conception is appealed to in order to frame the problem of fatalism: If God knows what I shall do before I do it, how can I avoid doing it and, therefore, how can I be held responsible for my deeds? In light of the contrast between the theocentric and anthropocentric models of cognition sketched in this chapter, however, this conception may also be used to understand the Kantian doctrine of

the ideality of time. The point here is simply that, insofar as it recognizes this atemporal conception of divine cognition as normative (as it must, if it is to preserve omniscience), it is transcendental realism that is led to conclude that time is not fully real, that objects and events only *appear* to be temporally successive. In other words, transcendental realism is confronted with a dilemma: it must either deny divine foreknowledge, which is philosophically difficult (though not unheard of), or deny the reality of time—that is, it must admit that occurrences merely *seem* to be successive but in reality they are not, which is to reduce experience to illusion.

The interpretation of transcendental idealism offered here provides a ready escape from this dilemma, thereby making it possible to preserve the empirical reality of time at the modest cost of its transcendental ideality. This is because considering time as an epistemic condition ensures its “objective reality” with respect to appearances, while also leaving conceptual space for a radically distinct atemporal perspective representing the God’s-eye view of things. Moreover, this advantage adheres only to a “one-world” understanding of this idealism, since what is required is that one and the same set of events be conceivable from these two radically distinct points of view. On a “two-world” reading, even of the “qualified” sort advocated by Van Cleve, this is obviously impossible, with the consequence that the dilemma remains as intractable as it is for transcendental realism in all its forms. Of course, as some philosophers have done, the transcendental realist may choose to bite the bullet and deny the reality of time. But it seems doubtful that Van Cleve and others who pursue his line of criticism would find that option attractive.



### B. Ameriks: Epistemology or Metaphysics?

The second basic objection to be considered here is a variant of the triviality charge, though it deserves a separate treatment because it has been raised by interpreters having more sympathy for transcendental idealism than either Guyer or Langton. A good representative of this approach is Karl Ameriks.<sup>52</sup> Ameriks’s objection is not that transcendental idealism on this interpretation is incoherent (though he may also believe that to be the case) but that it fails to do justice to Kant’s own understanding of his idealism. According to him, an epistemic interpretation simply ignores the ontological significance that Kant attaches to the transcendental distinction.<sup>53</sup> Thus, he claims, “On that [epistemic] reading there is still no reason to think the non-ideal has a greater ontological status than the ideal.”<sup>54</sup> But this, Ameriks thinks, is incompatible with Kant’s deepest philosophical commitments, which concern “the absolute

reality of things in themselves with substantive non-spatio-temporal characteristics.”<sup>55</sup>

Admittedly, there is much in Kant that suggests an ontological reading of the sort advocated by Ameriks. To begin with, it might be argued that “appearance talk” is only meaningful if it is contrasted with talk about things as they “really are.” Thus, even granting that the distinction is between two ways of considering things rather than between two kinds of thing, it would still seem that whatever can be said of things on the basis of the first way of considering them must have lesser ontological import than claims based on the second. In short, it seems that, under *any* interpretation, transcendental idealism must be seen as in some way incorporating the classical ontological contrast between appearance and reality.

Such a reading also seems to draw support from the fact that in the Inaugural Dissertation Kant explicitly contrasted sensible to intellectual cognition as a cognition of things as they appear to one of “things as they are” (Diss 2: 392; 384). Even though the “critical” Kant denied that we can have knowledge of the latter type, his continued adherence to the Dissertation’s doctrine of sensibility, and his equation of a consideration of things as they are in themselves with a consideration of things as some pure understanding might think them, certainly suggest that the ontological contrast of the earlier work is still operative.

Perhaps the strongest support for the ontological reading, however, appears to come from Kant’s moral philosophy, particularly his practical metaphysics of the supersensible. By affirming the primacy of practical reason or, what amounts to the same thing, denying knowledge in order to make room for faith (Bxxx), Kant is often taken as offering an entrée through practical reason to the very same ultimate reality that he had foreclosed to speculation. Accordingly, on this view we really are free, immaterial substances, and so forth, though we cannot demonstrate this theoretically.

Nevertheless, things are not that simple, since a straightforwardly ontological reading of the sort Ameriks (and many others) favor founders over the problem of empirical realism. As we have seen, once statements about things considered as they are in themselves are taken as claims about how they *really are*, it becomes difficult to avoid taking statements about appearances as claims about how they merely *seem to us* to be. And this, in turn, is hard to reconcile with any robust form of empirical realism. One obvious way of preserving this realism is Guyer’s proposal to jettison the idealism altogether. But this is to throw out the baby with the bath water. Short of that, however, there appears to be no solution available within the framework of Kant’s philosophy, save somehow deon-

tologizing the transcendental distinction. Whatever it may be, it *cannot be* a distinction between how things seem to be to beings like us and how they really are.

The conception of an epistemic condition was introduced precisely to resolve this problem. As already noted, the discursivity thesis looms large in this reconstruction, since it makes it possible to understand how discursive cognizers, such as ourselves, could have two radically distinct epistemic relations to objects, neither of which is *ontologically* privileged.<sup>56</sup> Ameriks questions, however, the compatibility of this approach with the “deeper” noumenalistic strains of Kant’s thought. Accordingly, it is this issue that we must now consider.

The matter is best approached in connection with the concept of freedom. On a traditional ontological reading, Kant is committed to the thesis that we really are (transcendentally) free agents, even though, when considered as phenomena, we are also causally determined parts of nature. Setting aside the question of how we could know this to be the case, given the impossibility of any theoretical knowledge of such freedom, the problem is to understand what this doctrine says about our phenomenal selves. Is it the case that we only seem to be causally determined, whereas we really are free? Or is it rather that our phenomenal selves really are determined and our noumenal selves really free? Neither alternative seems acceptable: the former because it undermines Kant’s empirical realism and the latter because it saddles him with an incoherent doctrine of two selves.

When approached in this way, it becomes clear that the heart of the problem is the underlying assumption that there is a “fact of matter” that needs to be adjudicated. On this assumption, the freedom, which, according to Kant’s moral theory, we are required to assume, must be viewed either as a real property of a separate noumenal self or as a property of our single self as it really is in its inner constitution. It is, however, just the assumption that there must be some standpoint-independent fact of the matter, which is implicit in any ontological reading of transcendental idealism, that is called into question by the interpretation advocated here. In fact, on this interpretation, such an understanding of transcendental idealism (like Van Cleve’s) is itself a form of transcendental realism.

Admittedly, this sounds extremely paradoxical, since we naturally tend to think that there must be some fact of the matter here. Either we are really free or we are not. One of these alternatives must be the case, even if we are not in a position to determine which one. Moreover, this holds whether we understand freedom in the Kantian sense as a non-natural causal power or in the popular contemporary compatibilist sense as a purely natural power for self-direction



and action. Paradoxical as it may be, however, this is precisely the conclusion to which Kant's transcendental idealism leads.

The view here ascribed to Kant can be made clearer by an appeal to Michael Dummett's conception of "warranted assertibility," which has been applied to the interpretation of Kant.<sup>57</sup> Although it is usually used in theoretical contexts to provide an anti-realist alternative to the traditional view of truth as correspondence to an independent reality or realm of facts, this conception may be extended to the practical context in which Kant discusses freedom. What is crucial here is Kant's contention that freedom is only assertible from a "practical point of view," that is, only in connection with our conception of ourselves as accountable moral agents. Clearly, Kant held that we must assume our freedom from that point of view.

It also seems clear, however, though it is more controversial, that he *did not* hold that this point of view provides access to some higher realm of being (the "really real"). The point is rather that from the practical point of view we are rationally authorized or warranted to assume our freedom, with the warrant stemming from the moral law as the law of pure practical reason. Correlatively, from the theoretical point of view, where the concern is with explanation rather than action, we are authorized, indeed required, to subject every event to the principle of causality as a condition of the possibility of its cognition. The argument for the former is beyond the scope of this study.<sup>58</sup> The argument for the latter will be considered in chapter 9. Here we need only note that Kant held both to be the case.

On this reading, then, transcendental idealism may be characterized as a doctrine of warranted assertibility relativized to a point of view. The basic idea is that each point of view (the theoretical and the practical) has its own set of norms on the basis of which assertions are justified and each involves considering its objects in a certain manner (as they appear and as they are thought of in themselves). But there is no context-independent truth or fact of the matter. Otherwise expressed, Kantian dualism is normative rather than ontological.<sup>59</sup>

Admittedly, Kant does speak on occasion, particularly in *Groundwork* III and the *Critique of Practical Reason*, of the idea of freedom or the consciousness of the moral law as giving us an entrée to an intelligible world or higher order of things, quite distinct from the sensible world of experience. Nevertheless, it is clear from the context that the superiority of the former to the latter is to be construed in axiological rather than ontological terms. What we supposedly become aware of is a higher set of values and a vocation [*Bestimmung*] to pursue them, not of our membership in some higher order of being. Similarly, in the

second *Critique*, Kant speaks of the primacy of practical reason in relation to the speculative (5: 119–21; 236–38). But this means only that our practical interest (in morality and the conditions of its possibility) is entitled to override our speculative interest in avoiding ungrounded claims and that the latter must therefore submit to the former. Once again, then, there is no thought of any access (cognitive or otherwise) to an ontologically superior order of being.

Still, such an “anodyne” analysis does not seem satisfactory. The question: “Are we *really* free”? keeps returning. And the answer: “Yes, but only from a practical point of view” appears to be either a dodge or a confusion, because we cannot help assuming that there must be *some* fact of the matter. Although this is true, Kant has an explanation for it. Moreover, this explanation is an essential, though generally overlooked, aspect of his transcendental idealism. It is to be found in the doctrine of transcendental illusion, which will be the centerpiece of the fourth part of this book.<sup>60</sup> For the present, it must suffice to note that the illusion is not that we are free, or, for that matter, that we are causally determined. It lies rather in the assumption that we must *really be* one or the other in some ontologically privileged, context-independent sense. Such an assumption is unavoidable for transcendental realism with its theocentric paradigm, but it is precisely what is called into question by Kant’s “Copernican revolution.”



Finally, if there is a general lesson to be learned from all of this, it is that transcendental idealism cannot be properly interpreted from the standpoint of transcendental realism, since it consists precisely in the denial of the validity (though not the naturalness) of that standpoint. Unfortunately, however, the ongoing debate concerning the nature and significance of this idealism attests to the fact that this lesson has not been learned.

## Chapter 3 The Thing in Itself and the Problem of Affection

Of all the criticisms that have been raised against Kant's philosophy, the most persistent concern the thing in itself, particularly the notorious claim that it, or the transcendental object, somehow "affects" the mind, thereby providing the content of experience, which is then shaped and articulated by the mind's sensible and conceptual forms. Any account of Kant's transcendental idealism must, therefore, include an analysis of this issue, and this is the task of the present chapter.

Building upon the discussion in the two previous chapters, it has a threefold aim and is divided into three parts. By considering some relevant texts and the analysis of Gerold Prauss, the first attempts to substantiate the claim that talk about things in themselves is to be understood as elliptical for talk about things considered as they are in themselves and that this way of considering things has a significant role to play in the philosophical activity of transcendental reflection. The second attempts to clarify the connection between the conception of things so considered and the related, though distinct, conceptions of the noumenon and the transcendental object. The third ad-

dresses the notorious problem of affection. It argues that the idea of a transcendental affection is both essential to Kant's position and does not, as is commonly assumed, involve a violation of critical principles.

# I. THE PROBLEM OF THE THING IN ITSELF: ANALYSIS AND RECONFIGURATION

The Kantian critique of transcendental realism considered in the preceding chapter establishes a significant polemical use for the concept of the thing in itself, by showing how the objects of human cognition are *not* to be regarded in a philosophical account, namely, as things that are *ausser uns* in the transcendental sense. Of itself, however, this does not seem to justify any positive use of the concept within the *Critique*. For the claim that transcendental realism treats mere appearances as if they were things in themselves no more entails that there are such things than the claim that a certain person acts "as if he were God" entails the existence of a deity.

At the same time, however, the analysis also suggests that there is something deeply misleading in this way of posing the problem, as if the worry were whether Kant has a warrant for assuming the existence of things in themselves distinct from the appearances with which we are directly acquainted. Indeed, the temptation to worry about the *existence* of things in themselves disappears once it is recognized that Kant is not primarily concerned with a separate class of entities, which, unlike appearances, would supposedly "be there" even if there were no finite cognizers).

As we have seen, the concern is rather with the familiar objects of human experience, *considered as they are in themselves*. Admittedly, the situation is complicated by the fact that Kant regards God and rational souls as non-sensible (intelligible) beings ontologically distinct from the sensible objects of human experience (and some of the passages cited below likewise allude to such distinct intelligible beings).<sup>1</sup> Nevertheless, it seems reasonable to conclude on the basis of the analysis offered in the first two chapters that those places in which Kant seems to suggest that the problem is *primarily* existential have to be interpreted in the light of a prior understanding of the transcendental distinction rather than, as is usually done, be made the basis for an understanding of this distinction.

The mandatory starting point for any serious, textually informed treatment of this topic is Gerold Prauss's analysis. Prauss points out that in discussing things in themselves Kant uses a number of different locutions, not all of which



are obviously equivalent.<sup>2</sup> First, there is the short form *Ding an sich* (and its variants *Sache*, *Gegenstand*, and *Object an sich*), which suggests that the referent is to a thing with a certain mode of existence (an *an sich* or independent existence). As Prauss notes, even though this locution is relatively rare in Kant, it is the preferred one in the literature.<sup>3</sup> By far Kant's most common locution is *Ding an sich selbst* (and its variants *Sache*, *Gegenstand*, and *Object an sich selbst*).<sup>4</sup> Although this difference is glossed over by the standard English translations, the longer form at least gestures toward the idea of a thing as it is in itself [*wie es an sich selbst ist*]. The crucial point, however, is that both the short and the long forms are to be construed as elliptical versions of the canonical "thing considered as it is in itself" [*Ding an sich selbst betrachtet*], where the *an sich selbst* functions adverbially to characterize *how* a thing is being considered rather than the kind of thing it is or the way in which it exists.<sup>5</sup> Prauss also suggests that, even though the long form is itself relatively rare, it occurs frequently enough to justify its canonical status and to indicate that the shorter forms, which Kant evidently favored for stylistic reasons, are to be understood in the light of it.<sup>6</sup>

Prauss's philological thesis is an important contribution to our understanding of the Kantian conception of the thing in itself and underscores the misguided nature of many of the standard criticisms noted in the first chapter. Nevertheless, it cannot be the whole story, since at least two significant problems remain, even if it is granted. First, there is an ambiguity in the notion of the consideration of something as it is in itself, which helps fuel the familiar misunderstandings and criticisms. On the one hand, it might be thought that to consider something as it is in itself is to take it as something that exists in itself, that is, as a *substantia noumenon*, equipped with intrinsic properties in the manner suggested by Langton. On the other hand, it might be taken to mean simply considering it as it is independently of its epistemic relation to human sensibility and its conditions. Let us call the former the ontological and the latter the epistemological sense of a consideration of something as it is in itself.

It is clear that to consider something as it is in itself in the former sense entails so considering it in the latter as well. For a noumenal substance's inherent nature certainly pertains to it independently of the way in which it affects us (if it affects us at all). But the converse does not hold, since something could be considered as it is in itself in the epistemological sense without also being thought as a separate *substantia noumenon*. It could, for example, be thought of (though not cognized) as a property of such a substance or as a relation between noumenal substances, both of which apply to things considered independently of their relation to human sensibility.<sup>7</sup> Moreover, if this point is overlooked (as it

usually is), it becomes all too easy to slide from considering something as it is in itself in the epistemological sense to considering it as existing *an sich selbst* in the ontological sense, with all of the problems this engenders.

Second, once this ambiguity is removed, the canonical formulation serves to reconfigure the problem, but not to resolve it. Appropriately reconfigured, the problem can be characterized more precisely as the need to provide some account of both the possibility and significance of a *consideration* of things as they are in themselves in the epistemological sense. To begin with, we might well ask what basis we have to assume that the things that supposedly appear to us under the subjective conditions of our sensibility are also anything in themselves apart from the way in which they appear. For all we can know, might they not be mere appearances? Furthermore, assuming that they are something in themselves and that we can somehow consider them in this way, what is the point of doing so, given our inability to cognize things so considered? In short, from what we have seen so far, there seems to be nothing to preclude a fully developed transcendental idealism from completely excising the consideration of things as they are in themselves.

In addressing this problem, it should prove useful to focus on some specific passages in which Kant appears to argue for the necessity of an appeal to things as they are in themselves or noumena (for the present these are treated as equivalent). The following are among the best known and widely discussed of these:

(1) This [the phenomena-noumena contrast] was the result of the entire Transcendental Aesthetic, and it also follows naturally from the concept of an appearance in general that something must correspond to it which is not in itself appearance, for appearance can be nothing for itself and outside our kind of representation; thus, if there is not to be a constant circle, the word "appearance" must already indicate a relation to something the immediate representation of which is, to be sure, sensible, but which in itself, without this constitution of our sensibility (on which the form of our intuition is grounded), must be something, i.e., an object independent of sensibility. [A251–52]

(2) In fact, if we view the objects of senses as mere appearances, as is fitting, then we thereby admit at the very same time that a thing in itself underlies them, although we are not acquainted with this thing as it may be constituted in itself, but only with its appearance, i.e., with the way in which our senses are affected by this unknown something. Therefore the understanding, just by the fact that it accepts appearances, also admits to the existence of things in themselves, and to that extent we can say that the representation of such beings as underlie the appearances, hence of mere intelligible beings, is not merely permitted but also unavoidable. [Pro 4: 314–15; 107–8]

(3) Yet the reservation must also be well noted, that even if we cannot cognize

these same objects as things in themselves, we at least must be able to think them as things in themselves. For otherwise there would follow the absurd proposition that there is an appearance without anything that appears. [Bxxvi-xxvii]

(4) Nevertheless, if we call certain objects, as appearances, beings of sense (*phaenomena*), because we distinguish the way in which we intuit them from their constitution in itself, then it already follows from our concept that to these we as it were oppose, as objects thought merely through the understanding, either the very same objects [*eben dieselben*] conceived in accordance with the latter constitution, even though we do not intuit it in them, or else other possible things, which are not objects of our senses at all, and call these beings of understanding (*noumena*). [B306]

(5) Now the doctrine of sensibility is at the same time the doctrine of the noumenon in the negative sense, i.e., of things that the understanding must think without this relation to our kind of intuition, thus not merely as appearances but as things in themselves. . . . [B307]

These and similar passages suggest two distinct lines of argument “in defense of the thing in itself,” each of which emphasizes one aspect of the Kantian conception and has its champions among interpreters. Neither, however, is fully satisfactory as it stands.

According to one line, which finds its strongest textual support in the second of the above-cited passages, the thought of things (existing) in themselves is not only admissible but also necessary because of the need to acknowledge a “cause” or “ground” of appearances. Although Kant here connects this inference with affection, which, as we shall see, is not precisely a causal relation, it is convenient to term this the “causal interpretation,” since it construes the relation between an appearance and a thing in itself to be that of an effect to its cause or ground.<sup>8</sup> An obvious problem with this interpretation is that it apparently requires that we take the appearance and the corresponding thing in itself as two distinct entities. Otherwise, how could we speak of a causal relation between them? But even if we ignore this problem, as well as the notorious difficulties associated with the notion of a “noumenal causality,” it is clear that this move cannot provide what is needed. For in order to take things as they are in themselves as the transcendental causes or grounds of appearances, we must obviously first be able to consider them as they are in themselves, which is the very point at issue.

The other passages suggest that Kant’s claim is semantic. On this reading, Kant is affirming a relation of logical implication between the *concept* of an appearance and the *concept* of a thing as it is in itself, rather than a causal relation between the entities subsumed under these concepts.<sup>9</sup> The basic thought is that

the expression *appearance* is parasitic upon, or at least correlative with, the expression *thing in itself*. Accordingly, the use of the former expression presupposes the legitimacy of the latter. Such an interpretation seems to find strong textual support in the contention in the first passage that the affirmation of “something which is not in itself appearance,” that is, a thing in itself, “follows from the concept of an appearance in general.” Similarly, in the same passage Kant contends that “the word ‘appearance’ ” must be taken as already indicating a relation to some such thing. Finally, in the third passage, Kant suggests that the denial of things in themselves would lead to the “absurd proposition that there is an appearance without anything that appears.”

Although it cannot be denied that these passages suggest such a line of argument, it appears to be no more successful than the previous one. To begin with, apart from the third passage, the expressions *appearance* and *thing in itself* again seem to refer to two distinct types of entity; the claim being that reference to entities of the former type (appearances) presupposes reference to those of the latter type (things in themselves). Indeed, so construed, it is reminiscent of the argument, often brought against linguistic versions of phenomenalism, that “sense-data language” cannot replace “material-object language,” because a reference to sense data is intelligible only if they are contrasted with material objects. As such, it is likewise not directly applicable to the transcendental distinction, understood as holding between two ways of considering one and the same thing.

Moreover, the attempt to modify the argument so as to make it relevant to this distinction does not initially seem plausible. To be sure, we can agree with Kant that it would be absurd to suggest that there can be an appearance without something that appears. As we have already seen, however, this does not license the conclusion that what appears is also something in itself distinct from what it appears to be. Why could not its appearance, suitably qualified to include ideal conditions, a multiplicity of perspectives and the like, be all that there is to it, so that there remains nothing left over to be considered “as it is in itself”?

The short answer is that such a position amounts to a Berkeleyian-style idealism or phenomenalism and, therefore, to a form of transcendental realism. Indeed, if Kant’s idealism is understood in this way (as it usually is), the problem of the thing in itself becomes intractable; for, as the above-cited texts suggest, one is then reduced either to a highly questionable causal inference or an obvious non sequitur. The situation looks quite different, however, when Kant’s conception of appearance or, more precisely, of things considered as they ap-

pear, is understood in the light of the interpretation of transcendental idealism sketched in the first two chapters.

As I have noted repeatedly, to consider things as they appear is to consider them in their relation to the sensible conditions under which they are given to the mind in intuition, that is, as *in uns* in the transcendental (but not the empirical) sense; just as to consider them as they are in themselves is to think them apart from all reference to these conditions, that is, as *ausser uns* in the transcendental sense. Clearly, however, in order to consider things in the former manner, it is necessary to distinguish the character that these things reveal as appearing (their spatiotemporal properties, and so forth) from the character that the same things are thought to possess when they are considered as they are in themselves, independently of the conditions under which they appear. Accordingly, the “absurdity” to which Kant alludes may be more appropriately characterized as considering something as it appears, or as appearing (in the transcendental sense), without, at the same time, contrasting this with the thought of how it may be in itself (in the same sense). In fact, these contrasting ways of considering an object are simply two sides of the same act of transcendental reflection, an act that Kant describes as “a duty from which no one can escape if he would judge anything about things *a priori*” (A263 / B319).

It must also be emphasized that the significance of this latter mode of considering things is not vitiated by the fact that it fails to provide any information about the nature of things so considered (not even the minimal information that they are substances containing unknowable intrinsic properties, which is thought by Langton to be necessary in order to ground Kantian humility). Indeed, paradoxically enough, this is the very source of its significance, which lies in its role within transcendental reflection.

First, the cognitive vacuity of a consideration of things as they are in themselves does not amount to incoherence. That would be the case only if the understanding could not even *think* things apart from the conditions of sensibility, which Kant repeatedly affirms that we can. As a result, there is nothing preventing a consideration of things as they are in themselves. It is only necessary to keep in mind that in considering things in this manner one must use pure categories (without any corresponding schemata) and that these can yield only analytic judgments about the *concept* of things so considered (as substance, cause, and so on).<sup>10</sup>

Second, such a consideration is necessary in order to avoid the “common assumption” of transcendental realism. Once again, the point is that in order to consider things as they appear (or qua subject to the conditions of sensibility),

which is itself necessary to explain the possibility of a discursive cognition of them, it is also necessary to contrast this manner of considering them with an alternative one. For unless things could at least be thought of as independent of the conditions of human sensibility, it would make no sense to require that they be considered in connection with these conditions. In other words, the significance of the consideration of things as they are in themselves (like that of the transcendental distinction itself) is directly methodological rather than metaphysical, even though (as we shall see later) it provides the only means for avoiding the metaphysical errors associated with transcendental realism.

## II. THE NOUMENON AND THE TRANSCENDENTAL OBJECT

We are now in a position to examine the concepts of the noumenon and the transcendental object, both of which are intimately connected with the concept of a thing considered as it is in itself. The former, which Kant first featured in his Inaugural Dissertation, we have already encountered on several occasions. Although like its opposite, that of a phenomenon, it appears to be an ontological concept, referring to a distinct kind of object, this object is characterized in epistemological terms as the correlate of a non-sensible and, therefore, purely intellectual cognition (just as a phenomenon is an object of a sensible cognition).<sup>11</sup>

The precise relation between this concept and that of a thing as it is in itself remains, however, a matter of some dispute. In particular, it seems natural for interpreters who affirm some version of a two-aspect reading to insist on a fairly sharp distinction between the two.<sup>12</sup> For whereas the former refers to the thing that appears as it is apart from the conditions under which it appears, the latter need not be so characterized. On the contrary, as is indicated by the examples of God and rational souls (both of which clearly count as noumena for Kant), a noumenon need not be the sort of entity that appears at all.

It is also noteworthy that Kant on occasion distinguishes sharply between their correlative concepts, namely, 'appearance' and 'phenomenon'. Thus, in his official definition of the former, he characterizes it as "[t]he undetermined object of an empirical intuition" (A20 / B34). If one emphasizes the term *undetermined*, which here means a lack of conceptual determination, then 'appearance' must be taken to refer to an object considered merely qua given in sensibility. This is to be contrasted with the phenomenon, understood as a sensible object that is brought under the categories. In short, a phenomenon is

a conceptually determined appearance.<sup>13</sup> But if this is the case, it seems that we must also draw a parallel distinction between the thing as it is in itself and the noumenon. The former is conceptually undetermined, since our thought of it is empty of real content, while the latter, as putative object of an intellectual intuition, is “conceptually” determined, though not for our discursive intellect.

Nevertheless, there are also many places in which Kant appears virtually to equate the two concepts.<sup>14</sup> Indeed, the basis for doing so can be traced to the Dissertation. As previously noted, having already arrived at the view that sensible cognition is conditioned by *a priori* forms imposed by the subject, Kant there concluded that such cognition yields knowledge of things as they appear (*uti apparant*); whereas a purely intellectual cognition (which he then deemed possible), since it is exempt from these conditions, must produce knowledge of things as they are (*sicuti sunt*) (Diss 2: 392; 384). Thus, according to the doctrine of the Dissertation, to cognize an object in a purely intellectual manner, that is, as a noumenon, is to cognize it as it is in itself, which holds whether or not we also have a sensibly conditioned cognition of the object as it appears. Moreover, even though in the *Critique* Kant rejected the view that we can know noumena, we have seen that he did not reject the concept of a noumenon. Rather, he sought to reinterpret it in such a way that it could be incorporated into his transcendental account. This is accomplished by giving it the function of a limiting or boundary concept (*Grenzbegriff*).

Kant develops this thesis in the chapter entitled “The Ground of the Distinction of all Objects in General into Phenomena and Noumena,” which provides a kind of overview of the results of the Analytic. At the heart of his position lies the contention that, despite its problematic status and merely negative use, the concept of a noumenon is “not invented arbitrarily” (A255 / B311). In denying its arbitrariness, Kant is presumably pointing out that the concept has a basis in transcendental reflection and is not, therefore, merely fictitious.<sup>15</sup> In other words, the “understanding” that must think noumena is the critical understanding, or, equivalently, the human understanding qua engaged in transcendental reflection. The point here is essentially the same as that previously made in connection with the concept of the thing considered as it is in itself. The critical understanding must think noumena because this concept is a correlate of the transcendental concept of appearance (or phenomenon), and is thus intimately connected with the doctrine of sensibility. In fact, it is just this connection with sensibility that enables it to function as a limiting concept. Its specific task is to “curb the pretensions of sensibility” (A255 / B311), which it

performs by referring, albeit in an indeterminate way, to a different manner of cognition (intellectual intuition), with respect to which the objects that appear to us as subject to sensible conditions would be cognized as they are in themselves, independently of these conditions.

In his initial treatment of the topic in the Dissertation, Kant used the "limitation of sensibility" brought about by the introduction of the concept of the noumenon to provide the conceptual space for a positive theory of the non-sensible. As already indicated, his goal was to clear the way for the cognition of an intelligible world by freeing our intellectual concepts from any "contamination" by our sensibly conditioned ones. In the *Critique*, by contrast, he notes that by limiting sensibility, which is accomplished by "calling things in themselves (not considered as appearances) *noumena*," the understanding also limits itself. It does so because it recognizes that it cannot cognize these noumena through any of the categories, but can merely think them "under the name of an unknown something" (A256 / B312).

In the first edition of the *Critique*, this "unknown something" becomes the "transcendental object." Nevertheless, precisely because it is an "unknown something," the latter must be sharply distinguished from the noumenon of the Dissertation, and at one point Kant explicitly does so. Thus, after noting that the object to which one refers appearance in general (the correlate of the transcendental concept of appearance) is "the transcendental object, i.e., the entirely undetermined thought of something in general," he remarks: "This cannot be entitled the noumenon; for I do not know anything of what it is in itself, and have no concept of it except merely that of the object of a sensible intuition in general, which is therefore the same for all appearances" (A253).


It seems strange for Kant to deny that the transcendental object is the noumenon on the grounds that we know "nothing of what it is in itself"; as if one *could* know what the noumenon is in itself! His point becomes understandable, however, if we take him to be denying that the transcendental object is equivalent to the noumenon of the Dissertation. *That* noumenon fell victim to the discursivity thesis, since Kant now claims that cognition of it would require intellectual intuition, of which we, as discursive knowers, are incapable. Since we lack such a capacity, that concept of a noumenon must be replaced by the concept of a transcendental object, construed as a mere indeterminate something. At the same time, however, Kant preserves a link with the terminology of the Dissertation by suggesting that the transcendental object may be referred to as a noumenon, provided this be correctly understood. Thus, in the Note to the Amphiboly of Concepts of Reflection he writes:



The understanding accordingly bounds sensibility without thereby expanding its own field, and in warning sensibility not to presume to reach for things in themselves but solely for appearances it thinks of an object in itself, but only as a transcendental object, which is the cause of appearance (thus not itself appearance), and that cannot be thought of either as magnitude or as reality or as substance, etc. (since these concepts always require sensible forms in which they determine an object). . . . If we want to call this object a noumenon because the representation of it is nothing sensible, we are free to do so. But since we cannot apply any of our concepts of the understanding to it, this representation still remains empty for us, and serves for nothing but to designate the boundaries of our sensible cognition and leave open a space that we can fill up neither through possible experience nor through the pure understanding. [A288–89 / B344–45]<sup>16</sup>

We shall return to this conception of the transcendental object as the “cause of appearance” in the next section of the chapter. Our immediate concern, however, is with the concept of the transcendental object itself, which proves to be particularly elusive, since Kant uses it in a number of different ways. In particular, the question concerns its relation to the thing as it is in itself. Much as is the case with the noumenon, there are places in which Kant seems to identify the transcendental object and the thing as it is in itself, including one (A366) where he does so explicitly, while there are others in which the two must be sharply distinguished.

The most interesting of the latter is in the A-Deduction, where Kant grants to the “pure concept of this transcendental object” the function of providing for “all our empirical concepts in general . . . relation to an object, i.e., objective reality” (A09).<sup>17</sup> Here Kant is concerned with the problem of objectivity or, more precisely, with the concept of an object corresponding to and distinct from our cognition (A104). In the preliminary portion of the Deduction, in which the concept of the transcendental object is introduced, the issue concerns what might be termed the “immanentization” of cognition, which is a direct consequence of Kant’s Copernican revolution. The basic problem is that we cannot, as it were, stand outside our representations in order to compare them with some transcendently real entity. Accordingly, such an object “must be thought only as something in general = X” (A104), which is later identified with the transcendental object. In this context, then, the concept functions as a kind of transcendental pointer, which serves to define the philosophical task by indicating that the commonsensical and transcendently realistic concern with the “real” nature of objects must be replaced by a critical analysis of the

conditions of the representation of an object. So construed, the term *transcendental object* obviously cannot be used in the plural, which is why Kant there describes it as being in all our cognitions “really always one and the same = X” (A109). As he later puts it, this transcendental object can serve only as a “correlate of the unity of apperception” (A250). 

Even at this early stage of the discussion, before we have examined Kant’s doctrine of apperception, it should be clear that such a correlate of apperception cannot be identified with the thing as it is in itself without doing violence to the critical nature of Kant’s project in the Deduction.<sup>18</sup> Nevertheless, in other places the relation between the transcendental object and the thing as it is itself seems much closer. In these places, some of which will be considered in the next section, the transcendental object serves as a transcendental ground of appearances rather than as a transcendental pointer. In other words, it is viewed as the (to us) unknown source of the sensible data. Since we can only intuit these data through our sensible forms, this transcendental ground may likewise be characterized only as a completely indeterminate, “something in general = X,” and thus again as a “transcendental object.”

Finally, albeit somewhat speculatively (since Kant himself never develops the point), it seems possible to specify an additional task, which can be performed only by the transcendental object and not by either the noumenon or the thing as it is in itself. Moreover, this task is necessitated by the understanding of the transcendental distinction affirmed here as involving two ways in which we can consider objects in transcendental reflection.

If, as is usually done, one approaches this distinction from a transcendentially realistic perspective, it is unavoidable to ask about the true nature of that which is supposed to be considered from these two points of view. But the attempt to answer this question on the basis of hints provided by the text appears to lead to a dead end. Thus, at one point Kant suggests that the object should be characterized as an appearance, on the grounds that the latter “always has two sides, one where the object is considered in itself . . . the other where the form of the intuition of this object is considered . . .” (A38 / B55). Although Kant’s explanation supports the thesis that it is the empirical object that is to be considered in two ways or from two points of view, it is certainly not correct to characterize what is so considered as an appearance or to suggest, as Kant does, that an appearance has two sides. It is rather the object that appears that has two sides, one of which is the way in which it appears (under the conditions of sensibility) and the other the way in which it is thought in itself (independently of these

conditions). Thus, what is being considered can be characterized neither as an appearance nor as a thing as it is in itself, since these correspond to the two ways of considering it.



Accordingly, Kant's best, perhaps his only, answer to this question is that it is the transcendental object that is considered from two points of view. In contrast to its role in the A-Deduction, however, its function here would be to dissolve a problem rather than to introduce one. As I have already noted, the problem it dissolves is one that arises when the transcendental distinct is approached from a transcendentially realistic point of view, which motivates the attempt to arrive at an absolute, context-independent conception of the objects of human cognition. It does so by indicating the vacuity of such a question; there simply is nothing more to be said about such an object than that it is a something in general = X.<sup>19</sup> As Kant succinctly puts it:

To the question: "What kind of constitution does a transcendental object have?" one cannot give an answer saying **what it is**, but one can answer that the **question** itself **is nothing**, because no object for the concept is given. [A478 / B506n]<sup>20</sup>

In order to understand why Kant would assign such distinct functions to the transcendental object, it is once again useful to consider that to which it is contrasted. The latter is the empirical object, and the contrast reflects Kant's distinction between the transcendental and the empirical. As he describes it at one point, "The difference between the transcendental and the empirical . . . belongs only to the critique of cognition and does not concern their relation to the object" (A56–57 / B81). In other words, transcendental cognition (the subject matter of the *Critique*) is not to be understood as providing knowledge of non-empirical (transcendental) objects. Its "object" is rather the conditions of the possibility of *a priori* knowledge of ordinary empirical objects or phenomena.



It is clear from this that the distinction between the empirical and the transcendental object, like that between things as they appear and as they are in themselves, is not between two ontologically distinct entities but between two perspectives from which ordinary empirical objects may be considered. Thus, talk about the transcendental object, like talk about things as they are in themselves, is to be construed adverbially as talk about empirical objects *considered transcendently*, that is, with respect to the conditions of their cognition.

Nevertheless, the transcendental object is not to be equated with the thing as it is in itself. In spite of their apparent identification at some points, they pertain to distinct levels of reflection, of which the transcendental-empirical dis-

inction is more fundamental. As the very terms indicate, the latter marks a contrast *between* a second-order philosophical consideration of objects and the conditions of their cognition (the business of the *Critique*) and a first-order investigation of them (the business of empirical science), while the distinction between things as they are in themselves and things as they appear is between two ways in which objects can be considered *within the transcendental standpoint itself*. Moreover, it is precisely because the concept of the transcendental object serves as an entrée to the transcendental standpoint that it can have such distinct functions. For these are just the distinct functions of the concept of an object considered transcendently.

The account offered so far was based on first-edition texts, although it includes references to passages that are retained in the second. The situation is further complicated, however, by the fact that Kant substantially revised the chapter on phenomena and noumena in the second edition. Moreover, in the revised version he not only dropped all references to the transcendental object but also introduced the distinction between a positive and a negative sense of the noumenon.<sup>21</sup> By the former he now understands “an *object* of a non-sensible intuition,” and by the latter “a thing insofar as it is not an object of our sensible intuition” (B307). The former is the rich concept of the noumenon from the Dissertation, and Kant’s point is once again that we cannot operate with this concept because we have no faculty of intellectual intuition and are not even able to conceive of its possibility. Thus he notes: “[T]hat which we call noumenon must be understood to be such only in a negative sense” (B309).

When we take the term in this sense, we do operate with the concept, since, as we have already seen, “the doctrine of sensibility is likewise the doctrine of the noumenon in the negative sense.” The thought expressed here is the by now familiar one that the concept of a non-sensible correlate of appearance is necessary to the very formulation of the Kantian theory. The noumenon in the negative sense is the second edition’s candidate for this required concept.

Nevertheless, this terminological shift does not reflect any significant doctrinal change. In fact, the distinction between the positive and the negative senses of the noumenon is only a more explicit reformulation of the first edition’s contrast between the noumenon and the transcendental object. For, as Kant now makes clear, the noumenon in the negative sense is not really a noumenon at all, except in the attenuated sense that it is something non-sensible. Moreover, since it refers merely to the non-sensible as such, it is completely indeterminate and, therefore, not distinguishable from the transcendental object.<sup>22</sup> As a re-

sult, even though Kant deleted references to the latter in those parts of the *Critique* that were rewritten for the second edition, he did not abandon the thought it expresses.

### III. AFFECTION

The Transcendental Aesthetic begins with the enigmatic and oft-cited words:

In whatever way and through whatever means a cognition may relate to objects, that through which it relates immediately to them, and at which all thought as a means is directed as an end, is intuition. This, however, takes place only insofar as the object is given to us; but this in turn, is possible only if it affects the mind in a certain way. [A19 / B33]

The reason this passage is so enigmatic lies in its deceptively innocent reference to the object that affects us. From Kant's time to our own, interpreters have continually returned to the question of the nature of this object. Indeed, this has been aptly termed "*die heikle Frage der Affektion*."<sup>23</sup> What gives this question its urgency is that, in spite of everything that Kant says about the unrecognizability of things as they are in themselves and the completely indeterminate nature of the concept of the transcendental object, there are numerous passages (one of which we have already noted) in which he refers to one or the other as affecting the mind or as the non-sensible "cause" or "ground" of appearances.<sup>24</sup> Moreover, this immediately poses a dilemma: on the one hand, there are important considerations suggesting that Kant must characterize the affecting object in some non-empirical manner, while, on the other, it has seemed to many that he cannot do so without violating a central tenet of the critical philosophy. This is, of course, just the problem posed by Jacobi in his famous and previously cited dictum that "*without* the presupposition [of the thing in itself] I cannot enter the [critical] system, and with that presupposition I cannot remain in it."<sup>25</sup>

Underlying this formulation of the problem is the understanding of affection as a kind of causal relation, which, as I have already indicated, is not entirely accurate. For affection, as Kant construes it, is clearly an epistemic rather than a causal relation, albeit one that is inextricably connected with the latter.<sup>26</sup> What makes the affection relation epistemic is that it holds between a discursive intelligence and what appears to it, that is, it involves a representation of an object, with this representation being determined in part by the mind's mode or manner of being affected (its forms of sensibility).



Nevertheless, it is also true that affection goes together with causation, since it involves an effect on the mind produced by the affecting object.<sup>27</sup> Moreover, this remains at the heart of the problem, even after affection is distinguished from causation. For, as already noted, Kant does speak of the transcendental object (or its equivalent) as affecting the mind, and he further seems to equate this with its serving as a non-sensible cause or ground of appearances.<sup>28</sup> Accordingly, it remains necessary to examine such assertions and the whole conception of affection in the light of the analysis given in the preceding section.

The natural starting point for such an examination is the account of Jacobi, which has set the agenda for virtually all subsequent discussions of the topic. Beginning with the uncontroversial premise that the Kantian theory of sensibility requires that the human mind be somehow affected by objects, Jacobi pointed out that there are only two possible candidates for the affecting object: an appearance and the transcendental object (which he equated with the thing in itself). Although Jacobi believed that Kant himself actually regarded the latter as the affecting object, his strategy was to show that neither can do the job. The former, he argued, cannot do it because it is defined by Kant as a mere representation in us (which is why we cannot enter into the critical system without the thing in itself); the latter cannot do it because of its uncognizability, which precludes the application to it of any of the categories, including causality (which is why we cannot remain in the system with it).<sup>29</sup>

In contrast to Guyer, who dismisses Jacobi's dilemma by rejecting the idealism from which it stems, idealistically oriented followers and "improvers" of Kant have tended to reject its first horn by denying that the *Critique* contains any doctrine of affection through things in themselves.<sup>30</sup> This ploy, however, encounters two difficulties: (1) It is apparently contradicted by passages in which it seems quite clear that Kant does recognize some sort of affection by non-sensible or intelligible objects (under various labels). (2) It does not explain how empirical affection, that is, affection through empirical objects or appearances, can provide the necessary ground of our representations. Vaihinger has usefully summarized the results of the post-Jacobi debate in the form of a trilemma:

1. Either one understands by the affecting objects the things in themselves; in which case one falls into the contradiction discovered by Jacobi, Aenesidemus and others that one must apply beyond experience the categories of substantiality and causality which are only supposed to have meaning and significance within experience.
2. Or one understands by affecting objects the objects in space; but since these are only appearances according to Kant, and thus our representations, one falls into the

contradiction that the same appearances, which we first have on the basis of affection, should be the source of that very affection.

3. Or one accepts a double affection, a transcendent affection through things in themselves and an empirical one through objects in space. In this case, however, one falls into the contradiction that a representation for the transcendental ego should afterwards serve as a thing in itself for the empirical ego, the affection of which produces in the ego, above and beyond that transcendental representation of the object, an empirical representation of the very same object.<sup>31</sup>

Although the theory of double affection, which was developed by Adickes,<sup>32</sup> continues to find some support among commentators,<sup>33</sup> it has been decisively repudiated from a number of different perspectives.<sup>34</sup> Among its many problems is that the attribution of such a theory to Kant, with its postulation of two distinct yet parallel activities, one of which is in principle unknowable, is incompatible with the critical nature of Kant's thought. According to Prauss, who develops his analysis with great subtlety, the whole problematic is grounded in a false "transcendent-metaphysical," as opposed to genuinely transcendental, conception of the thing as it is in itself. Prauss argues that once it is realized that this conception merely characterizes a manner in which empirical objects can be considered in transcendental reflection, it becomes obvious that nothing other than these same empirical objects affects us.<sup>35</sup> We thus return, albeit from a new perspective, to the old idealistic claim that the *Critique* admits only of an empirical affection; although Prauss readily acknowledges that this conception is itself not without difficulties.<sup>36</sup>

The problem of affection cannot, however, be dismissed quite so easily. In fact, the preceding analysis indicates that it, like the closely related problem regarding things as they are in themselves, has been misconstrued rather than, as Prauss suggests, that it is a pseudo-problem. As formulated by Vaihinger, the question concerns the metaphysical status of the affecting object: whether it is an appearance, a thing in itself, or perhaps both. Since this formulation is based on the assumption that the distinction between appearances and things in themselves is itself metaphysical in nature, once this assumption is repudiated, this way of stating the problem loses all meaning. But it does not follow that the problem itself disappears. It is still meaningful to ask whether Kant's statements about objects affecting the mind and producing sensations involve a reference to objects considered in their empirical character as appearances, or to these same objects considered as they are in themselves. If the former is the case, affection is to be construed in an empirical sense, and if the latter, in a transcendental sense.

Before addressing the problem in this new formulation, it is necessary to consider briefly Jacobi's contention that the very notion of an empirical affection is incompatible with the Kantian philosophy, because empirical objects are appearances and the latter are "mere representations in us." Although Jacobi is correct to deny that Kantian appearances can be the affecting objects, he is correct for the wrong reason, which makes his analysis deeply misleading. In fact, Jacobi's denial is based on a twofold confusion: it conflates Kantian appearances with Berkeleyan ideas, and it construes affection simply as a species of causation. Accordingly, if appearances are understood as mind-independent entities, considered as they appear in virtue of the subjective conditions of human sensibility, this whole line of objection dissolves. Kant can perfectly well speak of a causal (as opposed to an affective) relation between phenomena and the human mind, because at the empirical level the mind is itself part of the phenomenal world and subject to its conditions. Moreover, there are many places in which Kant speaks unproblematically in just this way.<sup>37</sup>

Nevertheless, the fact remains that a merely empirical account of affection, which reduces it to a causal relation between an affecting object and the mind or our sensory apparatus, cannot provide what Kant's transcendental account requires, namely, an *epistemic* relation between a discursive intelligence and the source of the matter or content of its sensible intuition. Thus, Jacobi is correct in ruling out an empirical affection as the source of this content, albeit, as noted above, for the wrong reason. The genuine reason (as suggested by Vaihinger) is conceptual in nature.

Let us recall that the Kantian theory of sensibility not only requires that something "affect" or be "given to" the mind; it also maintains that this something becomes part of the content of human cognition (the "matter" of empirical intuition) only as the result of being subjected to the *a priori* forms of human sensibility (space and time). It certainly follows that this something that affects the mind cannot be taken under its empirical description (as a spatiotemporal entity). To do so would involve assigning to an object, considered apart from its relation to human sensibility, precisely those features that, according to the theory, it only possesses in virtue of this relation. Consequently, the thought of such an object is, by its very nature, the thought of something non-sensible and hence "merely intelligible," a thing as it is in itself or a transcendental object, if you will.<sup>38</sup>

Against this, it is tempting to argue that even if one grants that the assertion "Something must affect the mind" has a transcendental status because it expresses a necessary (material) condition of the possibility of experience, this





does not require us to construe the expression *something* or its equivalents as doing anything other than referring indifferently to one or more of the members of the class of empirical objects. In other words (so the objection goes), all that has been established is the entirely general claim that some (empirical) object or other must affect the mind, if the mind is to have any experience. As a result, if this is the Kantian warrant for injecting the transcendental object (a “something in general = X”) into the transcendental story, then the whole account rests on a failure to distinguish between ‘something’, construed as having an indefinite reference, with ‘a something in general’, construed as a name or referring expression.

The problem with this line of criticism, which seems to be implicit in the endeavor to reject as unintelligible the very concept of a non-empirical affection,<sup>39</sup> is that it ignores one half of the transcendental account. As we have just seen, what is crucial to this account is that what is given to the mind as the result of its affection by external objects becomes part of the content of empirical cognition only by being subjected to the *a priori* forms of sensibility. But it follows from this that this “something” which affects the mind cannot, qua considered in this relation, be taken under an empirical description (as a spatiotemporal entity). For once again, to take it in this way requires assigning to the object, considered apart from its epistemic relation to human sensibility, precisely those features that, according to the theory, it only possesses in virtue of standing in this relation. To put it somewhat differently, if we think of Kant’s transcendental account of the conditions of discursive cognition as a “grand narrative,” then the indispensable role of material condition of this cognition must be assigned to something considered as it is in itself, apart from this epistemic relation and, therefore, as a merely transcendental object.

There is, however, nothing in this that commits Kant to the postulation of any super-empirical entities. On the contrary, no entities are assumed (in the account of affection) other than the spatiotemporal objects of human experience. The point is only that, insofar as these are to function in a transcendental account as material conditions of human cognition, they cannot, without contradiction, be taken under their empirical description. Consequently, we can reject the second horn of Jacobi’s dilemma, namely, that with the concept of the thing in itself (or transcendental object) we cannot remain in the Kantian philosophy. Indeed, we can remain in it quite comfortably, so long as we keep in mind that the concept has a legitimate metalevel function within the framework of transcendental reflection and, as such, does not bring with it any claim to cognize a super-sensible reality.

Finally, this analysis also provides the key to the interpretation of a number of the murkiest passages in the Kantian corpus. For illustrative purposes, it should suffice to consider two of these, one that does and one that does not refer to the transcendental object. Of those that do contain such a mention, the following is perhaps the most interesting:

The sensible faculty of intuition is really only a receptivity for being affected in a certain way with representations, whose relation to one another is a pure intuition of space and time . . . which, insofar as they are connected and determinable in these relations . . . according to laws of the unity of experience, are called **objects**. The non-sensible cause of these representations is entirely unknown to us . . . for such an object would have to be represented neither in space nor in time (as mere conditions of our sensible representation), without which conditions we cannot think any intuition. Meanwhile we can call the merely intelligible cause of appearances in general the transcendental object, merely so that we may have something corresponding to sensibility viewed as a receptivity. To this transcendental object we can ascribe the whole extent and connection of our possible perceptions, and can say that it is given in itself prior to all experience. But appearances are, in accordance with it, given not in themselves but only in this experience, because they are mere representations, which signify a real object only as perceptions, namely, when this perception connects up with all others in accordance with the rules of the unity of experience. Thus one can say: The real things of past time are given in the transcendental object of experience, but for me they are objects and real in past time only insofar as I represent to myself that, in accordance with empirical laws, . . . a regressive series of possible perceptions . . . leads to a time-series that has elapsed as the condition of the present time, which is then represented as real only in connection with a possible experience and not in itself. . . . [A494–95 / B522–23]




In this important and intriguing passage, Kant is addressing the question of how, on the basis of the doctrine of the ideality of time, it is possible to refer meaningfully to events in the distant past, before the advent of beings with our forms of sensibility. It falls naturally into two parts, each of which assigns a different role to the transcendental object. In the first part the context is the familiar one of the theory of sensibility. Here we note immediately the reference to the “non-sensible” and thus uncognizable cause of our representations, which is equated with the “merely intelligible cause of appearance in general.” As such, it is characterized as the “transcendental object” and assigned the function of providing “something corresponding to sensibility viewed as receptivity.”

One noteworthy feature of the first part of the passage is the shift from “appearance talk” to “representation talk,” which is a frequent and at times dis-

comforting feature of Kantian analysis. Nevertheless, insofar as ‘appearance’ is taken to refer to an object qua sensibly represented, we can allow Kant to talk indifferently about either the cause of representations or the cause of appearances. The really important point is the reference to “appearances in general.” This underscores the transcendental nature of the account by indicating that the concern is not with the cause of a given appearance or representation, which is always an empirical matter, but rather with the cause or ground of the “matter” of human cognition taken as a whole. Kant characterizes *this* cause as “non-sensible,” and hence uncognizable, precisely because it must not be represented as being in either space or time. Significantly, Kant does not say that such an object or cause cannot *be* in space or time, but merely that it may not be *represented* as in either. As a result, this prohibition must be understood methodologically; it stipulates how an object must be considered, if it is to function in a transcendental account as “something corresponding to sensibility viewed as receptivity.” As such, the prohibition does not bring with it any ontological assumptions about the real nature of things or about a super-sensible realm.

The second part of the passage is more obscure, in part because of its convoluted syntax, but mainly because Kant seems to expand the role of the transcendental object. Thus he allows us to ascribe to it “the whole extent and connection of our possible perceptions,” and even to say that “the real things of past time are given in the transcendental object of experience.” Expressions like these have led to the suggestion that the concept of the transcendental object functions as a kind of “conceptual repository” for our way of referring to the remote past or to distant regions of space.<sup>40</sup> On such an interpretation, it is effectively reduced to a high-order empirical concept, which refers to experience as a whole. The strategy behind such an interpretation is obviously to undercut the standard objections to Kant’s purported appeal to non-empirical objects. If it can be shown that Kant makes no such appeal, then these objections can be dismissed out of hand.

Although basically correct, such an analysis is somewhat misleading, since it ignores the fact that here, as elsewhere, the transcendental object is explicitly characterized in non-empirical terms. But this does not mean that Kant is here offering some kind of metaphysical explanation, which, in the manner of Berkeley’s appeal to the divine mind, allows him to “save” the reality of unperceived objects and events. The truth is just the opposite: once again the characterization of this ground as the transcendental object serves to repudiate any such metaphysical explanation, thereby enabling us to see that we must define

the reality of past events in terms of their connection with the present “in accordance with the rules of the unity of experience.” Accordingly, the point here is substantially the same as the one discussed in the preceding chapter in connection with the analysis of the difference between Kant’s idealism and phenomenalism or idealisms of the Berkeleian sort. Transcendental idealism is, at bottom, an idealism of epistemic conditions and, much as in the A-Deduction, the concept of the transcendental object functions as a reminder of this. 

The second passage comes from Kant’s reply to Eberhard. In pursuit of his avowed goal of demonstrating the superiority of Leibnizian rationalism to Kantian criticism, which involves showing the cognizability of things as they are in themselves, Eberhard had claimed supposedly, in opposition to the *Critique*, that things in themselves must be viewed as the source of the matter of sensibility. To this Kant replies:


Now that, of course, is the constant contention of the *Critique*, save that it posits this ground of the matter of sensory representations not once again in things, as objects of the senses, but in something super-sensible, which *grounds* the latter, and of which we can have no cognition. It says the objects as things in themselves *give* the matter to empirical intuition (they contain the ground of by which to determine the faculty of representation in accordance with its sensibility), but they *are* not the matter thereof. [UE 8: 215; 306–7]

Although this passage does not mention the transcendental object, it does appeal to the notion of a super-sensible ground, which is its functional equivalent.<sup>41</sup> The passage reflects the continuing debate concerning the Kantian theory of sensibility and its relation to the Leibnizian theory advocated by Eberhard. Kant himself poses the issue quite sharply when he notes that, on the critical view, sensibility is understood to be “only the mode in which we are affected by an object which in itself is entirely unknown to us” (UE 8: 219; 310). This is contrasted with the Leibnizian doctrine, according to which “we intuit things as they are in themselves” and, correlatively, “sensibility consists merely in the confusedness that is inseparable from such a multifaceted intuition” (UE 219; 310). The contrast drawn here is between their respective views on the relation between sensibility as a faculty and things as they are in themselves. Here, as in the *Critique*, Kant’s claim is that we cannot cognize the affecting object as it is in itself, because we can cognize an object only if it is given in intuition, and it can be given only in accordance with the mind’s *a priori* forms of sensibility (space and time). Since Leibniz does not acknowledge any such forms, he is committed to the transcendentially realistic doctrine that we sensibly appre-

hend things as they are in themselves. All of this is, of course, familiar from our previous discussion. It is, however, worth noting that for Kant, but not for Leibniz or Eberhard, what is given in intuition is sensible, regardless of its clearness or distinctness, whereas whatever cannot be so given but is merely thought, is non-sensible, intelligible, or, equivalently, super-sensible, again regardless of the clarity or distinctness with which it is thought.

This conception of the sensible and its opposite, here characterized as the “super-sensible,” underlies the distinction between the *matter* of sensible representation or empirical intuition and its *ground*, which Kant elsewhere terms ‘transcendental matter’.<sup>42</sup> Because he is confused about the nature of sensibility, Eberhard conflates these two concepts. It is also on the basis of this same confusion that he affirms, against Kant, the cognizability of things as they are in themselves, considered as the grounds of the matter of our sensible representations.<sup>43</sup> The point of Kant’s distinction between “ground” and “matter” is just to indicate the super-sensible nature of the former, in contrast to the sensible nature of the latter. The reason for characterizing the ground as “super-sensible” is precisely the same as in the previous passage, namely, its non-representability in space and time. Moreover, Kant naturally assigns this ground, as super-sensible, to “the objects as things in themselves,” rather than to “things, as objects of the senses.” Since ‘as’ (*als*) in both cases is obviously short for ‘considered as’, Kant can be taken to be merely affirming the by now familiar contention that the thought of an object as such a ground requires the consideration of it as it is in itself. Once again, Kant can say this because it is a merely analytic claim, based upon the concept of an object that is conceived of within a transcendental context as the ground of our representations.

The basic thrust of this analysis can be applied to a great many other passages, where Kant appears to be making illicit claims about the nature and function of things as they are in themselves or the transcendental object. So interpreted, these passages can be seen to accord with Kant’s critical and anti-metaphysical stance regarding the super-sensible in all its guises. Far from providing a metaphysical story about how the mind or noumenal self is somehow affected by a non-sensible entity, they merely stipulate how the affecting object must be conceived of in the transcendental account of affection required for the explication of the Kantian theory of sensibility. To be sure, these claims involve a use of the categories, especially causality, which is only to be expected, given the nature of the categories as concepts of an object in general. This does not, however, justify the frequently voiced criticism that Kant is guilty of an illicit application of the categories beyond the empirical realm. The function of the

categories in these transcendental contexts is purely logical, and does not carry with it any assumptions about their objective reality with respect to some empirically inaccessible realm of being. 

Accordingly, the proper answer to Jacobi, in contrast to the one proposed by Guyer, is to accept the first horn of the dilemma and reject the second. In other words, one can perfectly well remain within the critical philosophy with the concepts of the thing as it is in itself and the transcendental object, provided that they are properly, that is, non-metaphysically, understood. Moreover, we can see that Jacobi's dilemma, like so many supposedly devastating criticisms of Kant, reflects a transcendently realistic approach to the *Critique*. Underlying this approach is the assumption that Kant owes us, yet cannot provide, some ultimate metaphysical story about affection: a God's-eye account of what it is that *really* supplies the matter of cognition. Much the same view is also assumed by Vaihinger with his trilemma. But if the analysis of transcendental idealism presented in these three chapters is correct, it is precisely this assumption that must be rejected. In Wittgensteinian terms, Kant was not trying to say what is unsayable, but merely to define the boundaries of what can be said or asked. In order to do so, however, he had to introduce the "metalanguage" of transcendental philosophy. Thus, such expressions as "things as they are in themselves," "noumena," the "transcendental object," and their correlates are to be understood as technical terms within this metalanguage rather than as terms referring to transcendently real entities.<sup>44</sup>

## **Part Two Human Cognition and Its Conditions**

## Chapter 4 Discursivity and Judgment

I argued in the first part of this book that Kant's idealism, as an idealism of epistemic conditions, is inseparable from his analysis of the discursive nature of human cognition. I further suggested that this analysis is based on three bedrock epistemological assumptions: (1) that *cognition* of any kind requires that an object somehow be given (this applies even to the problematic intellectual or archetypal intuition); (2) that since a finite mind like ours is receptive rather than creative, its intuition must be sensible, resting on an affection by objects; and (3) that sensible intuition, of itself, is insufficient to yield cognition of objects and requires the cooperation of the spontaneity of the understanding.



Since for Kant at least the first two assumptions are relatively non-controversial, much of the weight falls on the third, which, as we have seen, constitutes a fundamental rejection of empiricist epistemologies.<sup>1</sup> Accordingly, this will be the main focus of the present chapter, which is divided into four parts. The first two are concerned, respectively, with Kant's conceptions of concepts and sensible intuitions (the "elements" of discursive cognition) and his account of the activity of



judgment in which these elements are united in the cognition of objects. In light of this, the third and fourth parts are concerned, respectively, with the analytic-synthetic distinction and the highly controversial conception of a synthetic *a priori* judgment. The chapter as a whole is intended as a prolegomenon to the analyses of the sensible and intellectual conditions of human cognition that follow in the next two chapters.

## I. DISCURSIVE COGNITION AND ITS ELEMENTS: CONCEPTS AND INTUITIONS

The *Critique of Pure Reason* is divided into two parts of unequal length and importance: The Transcendental Doctrine of Elements and The Transcendental Doctrine of Method. The former, which contains the bulk (but by no means all) of the philosophical significance of the work, is itself divided into two radically unequal portions: The Transcendental Aesthetic and The Transcendental Logic. The former is concerned with the contribution of sensibility and the latter with that of the understanding in the broad sense as the “higher cognitive faculty.” Since the latter encompasses both the understanding in the narrow sense and reason (*Verstand* and *Vernunft*),<sup>2</sup> it is likewise divided into two parts (The Transcendental Analytic and The Transcendental Dialectic). Although this division will prove to be of fundamental importance, it can be ignored for the present, since for an understanding of the Kantian conception of discursivity it is the nature and function of concepts that is central.

In the published version of his *Lectures on Logic*, Kant characterizes a concept (as opposed to an intuition) as “a universal representation, or a representation of what is common to several objects, hence a representation *insofar as it can be contained in various ones*.” Consequently, he points out, it is redundant to speak of universal or common concepts, as if concepts could be divided into universal, particular, and singular. “Concepts themselves,” Kant remarks, “cannot be so divided, but only their *use*” (JL 9: 91; 589). In the parallel account in the *Critique*, Kant notes that a concept, again in contrast to an intuition, refers to its object “mediately by means of a mark [*eines Merkmals*] which can be common to several things” (A320 / B377). In other words, because of its generality, a concept can refer to an object only by means of features that are also predicable of other objects falling under the same concept.

We saw in chapter 2 that it was precisely because of this inherent feature of concepts that Kant’s predecessors (both rationalist and empiricist), as well as the pre-critical Kant himself, tended to denigrate conceptual representation to

a second-class status, with the paradigm being a form of non-mediated, intuitive representation that grasps the object as it is in itself in its full concreteness. Although the critical Kant denies neither this feature of discursive cognition nor its inadequacy measured by some putative theocentric standard, he does deny the normativity of the latter standard for finite cognizers, such as ourselves. Rather than constituting a norm to which our cognition should conform, the theocentric model of a purely intuitive cognition is reconfigured as a limiting concept (providing the problematic idea of a mode of cognition with which our discursive kind is to be contrasted). Moreover, with this “paradigm shift,” discursive cognition attains for the first time its autonomy and normativity, as the form of cognition appropriate for finite rational beings.

An essential feature of this shift is the reinterpretation of concepts as rules, which Kant first announces in a casual way in the A-Deduction by stating that a concept, so far as its form is concerned, is always something universal that serves as a rule (A106).<sup>3</sup> As Béatrice Longuenesse has pointed out, however, there are two distinct senses in which concepts function as rules for Kant, which she correlates with the twofold meaning she finds in the term *concept* itself.<sup>4</sup>

On the one hand, concepts serve as rules of sensible synthesis, guiding the imaginative apprehension of particulars: for example, projecting the sides and back of a house seen from the front. To have the concept of a house is, among other things, to have a rule in the sense of a schema for organizing the sensory data given in perception.<sup>5</sup> On the other hand, concepts also serve as discursive rules affirming conceptual connections. Although the former sense of rule is certainly an important one for Kant, particularly, as we shall see, in the Transcendental Deduction, it is only the latter that is relevant to our immediate concerns. This is because it is the function of concepts as discursive rules that accounts for their role in judgment. To form the concept of body as a discursive representation is to think together the features of extension, impenetrability, figure, and so forth, as marks or components of the concept that are in some sense (to be explored later) necessarily connected in it. Correlatively, to apply this concept is to conceive of some actual or possible object or objects as falling under the general description provided by these marks. Since this is equivalent to forming a judgment, Kant claims that “the understanding can make no other use of these concepts than that of judging by means of them” (A68 / B93), and he characterizes concepts as “predicates of possible judgments” (A69 / B94).

Kant further distinguishes between pure (*a priori*) and empirical concepts and between the matter, or content, and form of a concept; but only the latter



distinction is directly relevant to our present concerns. By the *content* of an empirical concept Kant means the sensible features that are thought in it as its marks. These are derived from experience and correspond to the sensible properties of things. Thus, whatever is thought to be a body is also thought to be impenetrable and to have a certain size and shape and so on. By the *form* of a concept, Kant means its universality or generality, which is the same for all concepts and which accounts for their function as discursive rules. Simply having a set of sensible impressions that are associated with one another is not the same as having a concept, since the latter requires the thought of the applicability of these impressions to a set of possible objects.

This thought is not, however, itself derived from experience. Rather, it is produced by a series of “logical acts” of the understanding that Kant terms “comparison,” “reflection,” and “abstraction.” Taken together, these acts consist in the combining of the common, sensible features shared by diverse particulars into an “analytic unity,” while disregarding or abstracting from the differences (JL: 94–95; 592–93). Kant characterizes this whole process as “reflection” (*Reflexion, Überlegung*) (JL: 94; 591),<sup>6</sup> and a concept produced thereby a “reflected [*reflectirte*] representation” (JL 9: 91; 589).

Kant’s doctrine of empirical concepts and their formation is a complex and difficult topic, which is largely presupposed rather than articulated (much less argued for) in the *Critique*. Moreover, Kant’s discussions in the various versions of his *Lectures on Logic*, where he does deal with the issue, tend to be extremely cryptic and obscure. The basic problem, which cannot be pursued here, is that Kant’s official account of how we form such concepts, namely, by noting common features shared by diverse particulars and abstracting from the differences, seems to presuppose what it purports to explain. For how can one recognize such commonality without in a sense already having the concept?<sup>7</sup> Setting that issue aside, however, the crucial point for our present purposes is that concepts are not given or copied but made through an act of understanding.<sup>8</sup> As Kant puts it in the *Critique*, they are “grounded on the spontaneity of thinking” (A68 / B93). We shall see that this applies to pure or *a priori* as well as to empirical concepts.

In contrast to a concept, Kant defines an intuition in the *Lectures on Logic* as a “singular representation” (*repraesentatio singularis*) (JL 9: 91; 589). He repeats this in the *Critique*, adding that it is “immediately related to the object” (*bezieht sich unmittelbar auf den Gegenstand*) (A320 / B377).<sup>9</sup> Recognizing that the definition of ‘intuition’ as “singular representation” does not involve any reference to sensibility, Jaakko Hintikka has argued that only the singularity

criterion is essential and the immediacy criterion is a mere corollary.<sup>10</sup> This ignores, however, the presentational function of intuition, for it is in virtue of its “immediacy” that an intuition can present particulars to the mind and, therefore, serve as a *repraesentatio singularis*.

Nevertheless, a tension, if not outright contradiction, has often been noted between the official definition of ‘intuition’ as a singular representation and the discursivity thesis.<sup>11</sup> The problem lies in a point that has already been emphasized, namely, that according to this thesis, sensible intuition provides the mind only with the raw data for conceptualization, not with the determinate cognition of objects. As discursive, such cognition requires not merely that the data be given in intuition but also that they be taken under some general description or “recognized in a concept.” Only then can we speak of the “representation of an object.”<sup>12</sup>





The problem was already noted by J. S. Beck, who comments in a letter dated November 11, 1791:

The *Critique* calls “intuition” a representation that relates immediately to an object. But in fact, a representation does not become objective until it is subsumed under the categories. Since intuition similarly acquires its objective character only by means of the application of categories to it, I am in favor of leaving out that definition of “intuition” that refers to it as a representation relating to objects. I find in intuition nothing more than a manifold accompanied by consciousness (or by the *unique* “*I think*”), and determined by consciousness, a manifold in which there is no relation to an object. I would also like to reject the definition of “*concept*” as a representation mediately related to an object. Rather, I distinguish concepts from intuitions by the fact that they are thoroughly determinate whereas intuitions are not thoroughly determinate. For both intuitions and concepts acquire objectivity only after the activity of judgment subsumes them under the pure concepts of the understanding. [Br II: 311; 396]<sup>13</sup>

Kant’s only extant response to this query is contained in a marginal note attached to Beck’s letter. In it he remarks:

The fashioning [*Bestimmung*] of a concept, by means of intuition, into a cognition of the object is indeed the work of judgment; but the reference of intuition to an object in general is not. For the latter is merely the logical use of representation insofar as a representation is thought to belong to cognition. When, on the other hand, a singular [*einzelne*] representation is referred only to the subject, the use is aesthetic (feeling), in which case the representation cannot become an element of cognition [*Erkenntnisstück*]. [Br II: 311; 396–97]

As the latter part of the note indicates, Kant's response reflects the view of the recently published *Critique of Judgment*, in which he treats aesthetic judgments as non-cognitive because they are based on feeling. Consequently, he emphasizes the contrast between intuition and feeling. In so doing, however, he might seem to be conceding Beck's main point, since he acknowledges that, apart from being conceptualized in an act of judgment, intuitions do not really refer to or "represent" objects at all. 

Nevertheless, the note does provide an important indication of how Kant wished his definition of 'intuition' to be understood. The key lies in the reference of intuition to "an object in general." As the text indicates, this is essentially a matter of logical classification, that is to say, it is Kant's way of pointing out that intuitions, in contrast to feelings, *can* be brought under concepts in judgment and thereby be referred to particular objects. As W. H. Walsh has aptly put it, a Kantian sensible intuition is "proleptically" the representation of a particular.<sup>14</sup> To fulfill their representational function, intuitions must actually be brought under concepts, but their capacity to function in this way is sufficient to justify their logical classification. This also suggests, however, that it is necessary to distinguish between a determinate or conceptualized and an indeterminate intuition, only the former of which constitutes a *repraesentatio singularis*. Moreover, we shall see that this applies to both pure and empirical intuition. 

Unfortunately, this does not exhaust the complexity or, perhaps better, the ambiguity inherent in the Kantian conception of intuition. In fact, it applies to only one of three senses in which Kant uses the term: that in which it refers to a particular kind of representation or mental content. In addition to this more or less official sense of 'intuition', Kant also uses the term to refer both to the *object* represented by such a content (the *intuited*) and to the *act* of directly representing a particular (the *intuiting*). In short, it is necessary to distinguish between a mental content, an object, and an act sense of 'intuition'.<sup>15</sup> Moreover, while it is generally clear from the context when the term is being used in the third sense, it is often difficult to determine whether it is being used in the first or second sense or, indeed, whether Kant conflates the two senses. We shall see that a good deal rides on these questions; but before we are in a position to deal with this issue, we must consider Kant's general account of judgment.

## II. KANT'S CONCEPTION OF JUDGMENT

As I have already indicated, discursive cognition is judgmental. It is in and through judgments that we apply concepts to given data, while concepts them-

selves are characterized as “predicates of possible judgments.” Kant makes this explicit when he states, “We can . . . trace all actions of the understanding back to judgments, so that the understanding in general can be represented as a faculty for judging” (*ein Vermögen zur urtheilen*) (A69 / B94). One of the main problems, however, confronting any interpretation of Kant’s conception of judgment is that he defines ‘judgment’, meaning both the act (judging) and the product (the judgment), in a variety of ways. In the *Jäsche Logic*, for example, he defines a judgment in general as “the representation of the unity of the consciousness of various representations, or the representation of their relation insofar as they constitute a concept” (JL 9: 101; 597). In the *Wiener Logic*, by contrast, he writes more expansively:

A judgment is *generaliter* the representation of the unity in a relation of many cognitions. A judgment is the representation of the way that concepts belong to one consciousness universally[,] objectively. If one thinks two representations as they are combined together and together constitute one cognition, this is a judgment. In every judgment, then, there is a certain relation of different representations insofar as they belong to *one* cognition. E.g., I say that man is not immortal. In this cognition I think the concept of being mortal through the concept of man, and it thereby happens that this cognition, which constitutes the unity of two different representations, becomes a judgment. [WL 24: 928; 369]

Although in both places Kant indicates that an act of judgment involves the unification of distinct representations in a concept that is correlated with a unity in the consciousness of these representations, the second goes well beyond the first by suggesting that this unification pertains to consciousness *universally* and, therefore, *objectively*. In other words, it indicates that objectivity or “objective validity” is correlated with a certain unification in consciousness (one that purports to be universal) and is an inherent feature of judgment as such.<sup>16</sup> Moreover, a substantially similar contrast is to be found in the two accounts of judgment in the *Critique*. The first, which corresponds to (though it goes considerably beyond) that of the *Jäsche Logic*, is located in the discussion of the Logical Employment of the Understanding. As such, it amounts to an analysis of judgment from the standpoint of general logic (which abstracts from the properly epistemological question of objective validity). The second, which develops the core idea expressed in the *Wiener Logic*, is located in §19 of the B-Deduction. Consequently, it considers judgment from the standpoint of transcendental logic (where the question of validity is the central focus).

Because of this difference, it is sometimes claimed that these texts embody

two distinct, even incompatible, conceptions of judgment.<sup>17</sup> Nevertheless, if one keeps in mind the connection of the former with general logic and the latter with transcendental logic, it is possible to see them as concerned with different aspects of a single conception that is considered from two points of view.<sup>18</sup> To judge is *both* to unify representations by combining them in a concept (producing an analytic unity) *and* to relate these same representations to an object in a manner that purports to be valid with respect to the object. In order to clarify this important point, it will be necessary to consider briefly each of these accounts. This should then put us in a position to consider the nature and significance of Kant's distinction between analytic and synthetic judgments, which belongs entirely to transcendental logic.


### A. Concepts and Judgment:

#### The Initial Account

Kant's concern in the first of these accounts in the *Critique* is to make explicit the identification of discursive cognition and judgment. Every judgment for Kant involves an act of conceptualization, and vice versa.<sup>19</sup> Since Kant's conception of concepts commits him to the doctrine that "a concept is . . . never immediately related to an object, but is always related to some other representation of it (whether that be an intuition or itself already a concept)," he proceeds to define judgment as "the mediate cognition of an object, hence the representation of a representation of it" (A68/B93). Immediately after this definition, Kant provides a capsule account of his theory of judgment. Because of its brevity and importance, it is worth quoting in full.

In every judgment there is a concept that holds of many, and that among this many also comprehends a given representation, which is then related immediately to the object. So in the judgment, e.g., "**All bodies are divisible**," the concept of the divisible is related to various other concepts; among these, however, it is here particularly related to the concept of body, and this in turn is related to certain intuitions [or appearances]<sup>20</sup> that come before us. These objects are therefore mediately represented by the concept of divisibility. All judgments are accordingly functions of unity among our representations, since instead of an immediate representation a higher one, which comprehends this and other representations under itself, is used for the cognition of the object, and many possible cognitions are thereby drawn together into one. [A68–69/B93–94]

By indicating that judgment involves the relation of representations to an object, this goes beyond the definition in the *Jäsche Logic*. This should not surprise us, however, since it follows from the very nature of judgment as a cogni-

tion, which, as such, always purports to relate to  object. Once again, it may be described as a matter of logical classification. Consequently, it does not remove it from the domain of general logic, which for Kant encompasses the analysis of the discursive activities of the mind, considered as such.<sup>21</sup> This would be the case only if the normative question of the validity of such cognitions and its transcendental grounds were introduced, which does not occur at this point.

We also see from Kant's example that the judgment involves two concepts, 'body' and 'divisibility', which are related both to each other and to the object judged about, that is, to the complete set of  $x$ 's thought under the general description contained in the concept 'body'. Of these, the subject concept, 'body', stands in the more direct, though still not immediate, relation to the object. It does not relate to the object immediately, because no concept can do that, but rather to an immediate representation of it. Moreover, since the latter is, by definition, an intuition, the subject concept in Kant's illustration refers directly to an intuition, and only mediately through it to the object. As we have already seen, the intuition provides the sensible content for the judgment, while the concept provides the discursive rule by means of which this content is thought. It is precisely by determining this content that the concept is brought into relation with the object. That is why Kant characterizes the relation between concept and object as mediate.

Finally, the judgment asserts that the object so determined (the subject of the judgment) is also thought through the predicate 'divisibility'. This is a second determination or conceptualization of the object, one that is mediated by the first. It is this second determination to which Kant refers when he claims that in a judgment "many possible cognitions are . . . drawn together into one." The collection or unification effected by this particular judgment is of the  $x$ 's thought through the concept 'body' with other  $x$ 's that may be thought through the concept 'divisibility'. Kant's claim that "all judgments are functions of unity among our representations" is intended to underscore the point that every judgment involves a unification of representations under a concept, that is, an act of conceptualization, which, at the same time, involves the relation of these representations to an object. As we shall see in more detail in chapter 6, the term *function* here must be taken in the Aristotelian sense as equivalent to 'task' or 'work'. Kant is thus saying that the essential task of every act of judgment is to produce such an object-related unity of representations under a concept.

More detailed accounts of essentially this same conception of judgment are



to be found in many of Kant's *Reflexion*.<sup>22</sup> Although these accounts are generally intended as introductions to the distinction between analytic and synthetic judgments, the treatment of the generic features of judgment can be considered independently of that issue. One of the most important of these is referred to by Paton.<sup>22</sup> In the relevant portion of this *Reflexion*, Kant writes:

Every object is known only through predicates which we think or assert of it. Before this, any representations that may be found in us are to be regarded only as material for cognition, not as themselves cognitions. An object, therefore, is only a something in general which we think to ourselves through certain predicates which constitute its concept. Every judgment, therefore, contains two predicates which we compare with one another. One of these, which constitutes the given cognition of the object, is called the logical subject; the other, which is compared with it, is called the predicate. When I say 'a body is divisible' this means that something *x*, which I cognize through the predicates that together constitute a concept of body, I also think through the predicate of divisibility. [R 4634 17: 616–17]<sup>23</sup>

The first two sentences reiterate the previously made point about unconceptualized representations. Of more immediate significance, however, is the fact that Kant infers from this that every judgment must have two predicates. This claim cannot simply be accepted as it stands, since it applies only to categorical judgments; hypothetical and disjunctive judgments can have many more than two predicates. Nevertheless, since Kant regards these as logical compounds of categorical judgments, this can be ignored for the present.<sup>24</sup> The crucial point is that when Kant characterizes concepts as "predicates of possible judgments" he is not limiting their function to that of logical or grammatical predicate. His point is rather that predicates, that is, concepts, function to determine the very content to be judged about, which they do by providing a general description under which this content can be thought. Insofar as a concept fulfills this function, it is regarded as a "real" rather than a merely "logical" predicate. Such a predicate is also called a "determination" (*Bestimmung*).<sup>25</sup>

In the judgment under consideration, the logical subject, 'body', functions as a real predicate. In Kant's own terms, it "constitutes the given cognition of the object," which means that it provides the initial description under which the subject *x* is to be taken in the judgment. Since the judgment is analytic, 'divisibility' is only a logical predicate, which, as such, does not add any further determinations to the subject beyond those already contained in the thought of it as a body. Leaving aside for the present the whole question of analyticity, we see that the judgment "compares" these predicates with one another and asserts

that they pertain to an identical  $x$ . It thus asserts that the same (or some, or every)  $x$  that is thought through the predicate ‘body’ is also thought through the predicate ‘divisibility’. This is the basic Kantian schema for judgments of the categorical form, whether analytic or synthetic, which, as such, pertains to general logic.

## B. Judgment and Objectivity:

### The Second Account

As befits its location in the text, the objectivity of judgment is the focal point of the discussion of judgment in §19 of the B-Deduction. Kant is here concerned with the explication of the distinction, first made in §18, between an “objective unity” of self-consciousness, which presumably involves the categories, and a “subjective unity,” which is a product of the reproductive capacity of the imagination. He begins by criticizing logicians who define a judgment as the “representation of a relation between two concepts.” Although Kant notes in passing that the definition is inadequate, because it applies only to categorical judgments, his real complaint is that it does not specify in what this relation consists. In an effort to answer this question Kant writes:

I find that a judgment is nothing other than the way to bring given cognitions to the **objective** unity of apperception. That is the aim of the copula **is** in them: to distinguish the objective unity of given representations from the subjective. [B141-42]



The distinguishing characteristic of a judgment thus lies in its objectivity. It is an “objective unity” and, as such, is correlated with the objective unity of apperception. In view of Kant’s definition of the objective or transcendental unity of apperception, this means that every judgment involves “that unity through which all of the manifold given in an intuition is united in a concept of the object” (B139). For the purposes of this preliminary discussion, this can be taken to mean that every judgment involves a synthesis or unification of representations in consciousness, whereby the representations are conceptualized so as to be referred or related to an object.



So far this tells us nothing that could not be gleaned from the previous analysis. But Kant proceeds to remark that a judgment can be described as “a relation [of representations] that is objectively valid, and that is sufficiently distinguished from the relation of these same representations in which there would be only subjective validity, e.g., in accordance with laws of association” (B142). This effectively makes objective validity into a constitutive feature of judgment, rather than a value assigned to some judgments.

If this claim is to make any sense, objective validity cannot be equated with truth; otherwise, Kant would be committed to the absurdity that every judgment is true, simply in virtue of being a judgment. Thus, it seems reasonable to follow Prauss, who suggests that ‘objective validity’ here means simply the capacity to be true or false.<sup>26</sup> On this interpretation, Kant’s claim that every judgment is objectively valid is equivalent to the thesis that every judgment makes a claim to truth and, therefore, has a truth value. This clearly holds of all genuinely empirical judgments, though it raises certain problems about metaphysical judgments that need not concern us here.<sup>27</sup> It does not, however, hold of a merely imaginative or associative unification of representations, such as the association of heat with the thought of the sun. The latter is simply an event in one’s mental history. As such, it can be neither true nor false, which is not to say that one cannot form true or false judgments about it.

Although we shall return to the contrast between an objective or judgmental and a purely subjective imaginative or associative unity in chapter 7, it may be useful to dwell on it a bit longer here for the lessons it provides concerning Kant’s conception of judgment. Of particular importance in this regard is Kant’s account of the relation between the concepts united in a judgment, about which nothing has yet been said. As Longuenesse has pointed out, this relation is one of subordination, which Kant contrasts with the relation of coordination that holds of representations united by association.<sup>28</sup>

In the endeavor to clarify this important point, it is convenient to begin with the latter relation. Here Hume’s principles of association, resemblance, contiguity in space and time, and cause and effect (at least as a natural relation), are good examples. In each case, the associated representations (“perceptions” for Hume) may be said to be “coordinated” in the sense that one of the associated items automatically triggers the idea of another, apart from any discursive activity. That is why the products of such association have no truth value. By contrast, the subordination of concepts in a judgment amounts to a *conceptual* dependence. Thus, in the judgment “All bodies are divisible,” the predicate ‘divisibility’ is subordinated to the subject concept ‘body’ and, as such, is predicable of whatever the latter is (as well as of other objects of which the latter is not).

Since the relation of conceptual dependence or subordination holds for *any* consciousness (or, as Kant puts it in the *Prolegomena*, “consciousness in general”), this accounts for both the universality of the connection and its objective validity. At least it accounts for the latter, if, as Kant maintains, “the objective validity of a judgment of experience signifies nothing other than its necessary



universal validity" (Pro 4: 298; 92). Rather than constituting a regression to an empirical idealism or some kind of consensus view of truth, this must be seen as an expression of Kant's rejection of transcendental realism. As we have seen, the essential feature of this rejection is the replacement of the traditional theocentric, God's-eye view of cognition with one based on the conditions of human cognition. From this standpoint, objective validity must be understood as the "necessary universal validity" for beings with our forms of cognition. There simply is no higher norm to which one can appeal. We shall see in chapter 7 that this also accounts for the element of necessity, which, according to Kant, is built into even empirical judgments. Our immediate concern, however, is with the connection between Kant's conception of judgment and his controversial analytic-synthetic distinction.



### III. THE ANALYTIC-SYNTHETIC DISTINCTION

It is clear from Kant's discussions of the topic in the *Reflexionen*, various versions of his *Lectures on Logic*, and his response to Eberhard that the distinction between analytic and synthetic judgments is deeply rooted in his conception of judgment and, therefore, of the discursive nature of human cognition. Unfortunately, however, none of this is apparent from Kant's account of the distinction in the Introduction to the *Critique*. Moreover, this is one of the main reasons the distinction has been so frequently misunderstood and subjected to so much misguided criticism.

The Introduction contains two distinct, but purportedly equivalent, formulations of this distinction. According to the first, analytic judgments are those in which "the predicate B belongs to the subject A as something that is (covertly) contained in this concept A," or, equivalently, those in which the connection of the predicate with the subject is "thought through identity." Synthetic judgments, by contrast, are those in which "B lies entirely outside the concept A, though to be sure it stands in connection with it." The connection between subject and predicate in such judgments is thus said to be "thought without identity" (A6 / B10-A7 / B11). According to the second formulation, the distinction is between merely explicative (analytic) and ampliative (synthetic) judgments. The former "do not add anything [through the predicate] to the concept of the subject, but only break it up by means of analysis into its component concepts, which were already thought in it (though confusedly)." The latter, by contrast, "add to the concept of the subject a predicate that was not thought in it at all, and could not have been extracted from it through any

analysis" (A7 / B11). Only much later does Kant make explicit what is implicit in his entire discussion, namely, that the law of contradiction is the principle of all analytic judgments (A151 / B191). In the *Prolegomena*, Kant follows the second formulation, but he adds that the distinction concerns the *content* of judgments rather than their origin or logical form. Moreover, he states explicitly that analytic judgments depend wholly on the law of contradiction, and that this is a basic point of contrast with synthetic judgments (Pro 4: 266–67; 62).

The first formulation, which is the one usually cited, is particularly subject to misinterpretation, because it suggests that the distinction is a logical one, concerning the relation between the subject and predicate concepts in a judgment (whether or not one is included in the other). Indeed, this seems to support Eberhard's contention, summarily dismissed by Kant, that the analytic-synthetic distinction is equivalent to the distinction between identical and nonidentical judgments.<sup>29</sup> It also gives rise to the frequently stated objection that the distinction applies only to categorical judgments and thus cannot have the universality Kant claims for it. The major problem, however, is that this formulation provides no hint as to how syntheticity is to be understood (except as the negation of analyticity), or why Kant should insist in the *Prolegomena* that the distinction concerns the content rather than the logical form of judgments.

Beyond this, there are the familiar difficulties concerning how one determines whether one concept is "contained" in another. As Lewis White Beck has pointed out, Kant seems to recognize two distinct criteria for deciding such questions. He calls one "phenomenological" and the other "logical."<sup>30</sup> According to the former, the question whether one concept is contained in another is resolved by introspection: we reflect on what is "actually thought" in a given concept. According to the latter, the question is resolved by considering whether the contradictory of the original judgment is self-contradictory. If it is, then the original judgment is analytic and its truth can be determined in accordance with the principle of non-contradiction; if not, it is synthetic.

An obvious problem here is that these two criteria do not always produce the same result; it seems possible that a judgment could be analytic under one criterion and synthetic under the other. At best, the first version does nothing to dispel such a possibility. Moreover, the phenomenological criterion itself seems to be an unreliable guide, since it leaves unexplained how, in any given instance, one can determine whether the failure to find one concept contained in another is due to the syntheticity of the judgment or the limited insight of the person making the judgment. In short, it leaves open the possibility that any and every *apparently* synthetic judgment is covertly analytic. Unfortunately, the

logical criterion fares no better. The problem with it is that, except for manifest tautologies, it cannot be applied without appealing to “phenomenological” considerations, that is, to meanings. How, after all, could one determine whether the contradictory of a given judgment is self-contradictory without appealing to the meanings of the terms and, therefore, without determining whether the one concept is “contained” in the other?<sup>31</sup>

Although it hardly resolves all of these difficulties, the second version is superior to the first because in it the notion of a synthetic judgment, the real focus of Kant’s concern, “wears the trousers.”<sup>32</sup> We learn that a synthetic judgment is one through which we extend, rather than merely clarify, our knowledge. This both indicates that the two species of judgment differ in their epistemic functions and suggests why Kant insists in the *Prolegomena* that the distinction concerns the content of the judgments. In addition, it puts to rest the objection that the distinction is only relevant to judgments of the subject-predicate form.

Nevertheless, it does not tell us in what sense and by what means we extend our knowledge through synthetic judgments, and it retains much of the psychological or subjectivistic flavor that is suggested by the appeal to the phenomenological criterion in the first.<sup>33</sup> In order to understand how Kant deals with these problems, it is useful to consider the brief discussion in the *Jäsche Logic*. Kant there presents the analytic-synthetic distinction as a contrast between a “formal” and a “material” extension of knowledge. Analytic judgments, he tells us, extend our knowledge in the former and synthetic judgments in the latter sense (JL 9: III; 606–7).<sup>34</sup>

Analytic judgments provide a formal extension of knowledge by clarifying or explicating what is already implicit in a concept. This involves the uncovering of implications of which one may not have been previously aware, but which are derivable by strictly logical means from a given concept. Once again, Kant takes “All bodies are extended” as an example of an analytic judgment, which he renders schematically as “To everything  $x$  to which the concept of body ( $a + b$ ) belongs, belongs also *extension* ( $b$ )” (JL 9: III, §36; 607). This is the basic schema for an analytic judgment. It shows that in such judgments the predicate ( $b$ ) is related to the object  $x$  by virtue of the fact that it is already contained (as a mark) in the concept of the subject. Analytic judgments are, therefore, “about” an object; they have a logical subject and, as Kant’s example shows, they can also have a real subject. Nevertheless, since the truth or falsity of the judgment is determined merely by analyzing the concept of the subject, the reference to the object  $x$  is otiose.<sup>35</sup> That is why it is possible to form analytic judgments



about nonexistent, even impossible objects, and why all analytic judgments are *a priori*.


In his response to Eberhard, who pressed him on the distinction from a Leibnizian point of view, Kant supplements this account by introducing what amounts to a distinction between immediately and mediately analytic judgments.<sup>36</sup> “All bodies are extended” is immediately analytic, because ‘extension’ (together with ‘figure’, ‘impenetrability’, and so forth.) is a mark of the concept ‘body’. In the scholastic terminology introduced by Eberhard, these marks are parts of the “logical essence” of the concept. By contrast, “All bodies are divisible” is mediately analytic, because ‘divisibility’ is not itself part of the concept (logical essence) of body, but rather of one of its constituent concepts (*extension*). In other words, it is a mark of a mark. Although this implies that the judgment rests on an inference, and in that sense extends our knowledge, Kant insists that this does not amount to a difference in kind, since in each case the predicate is derived from the concept of the subject by a process of analysis.

This should suffice to show that Kant’s conception of analyticity is of a piece with the discursivity thesis. As the preceding account indicates, it rests upon his conception of a concept as a set of marks (themselves concepts), which are thought together in an “analytic unity,” and which can serve as a ground for the cognition of objects. These marks collectively constitute the intension of a concept. One concept is contained in another, just in case it is either a mark of the concept or a mark of one of its marks. In either case, it is subordinated to the concept in which it serves as a mark, which is precisely the relation that is brought out in an (affirmative) analytic judgment. Thus, unlike most contemporary conceptions of analyticity, Kant’s is thoroughly intensional. As Beck points out, it rests upon the doctrine of the fixity of a concept, that is, on the thesis that the marks of a concept can be sufficiently determined (even without an explicit definition) for the purpose of analysis.<sup>37</sup> The notorious difficulties that arise concerning analytic judgments involving empirical concepts, such as water, which we need not consider here, all stem from the difficulty of sufficiently determining such concepts.<sup>38</sup>

A synthetic judgment, by contrast, extends our knowledge in a “material” sense. Kant’s example is “All bodies have attraction,” which he renders schematically as “To everything  $x$ , to which the concept of body ( $a + b$ ) belongs, belongs also *attraction* ( $c$ )” (JL 9: 111; 607). Like its analytic counterpart, this judgment asserts a connection between the predicate ( $c$ ) and the subject ( $x$ ), which is thought through the concept ( $a + b$ ). But unlike its analytic counterpart, it asserts this independently of any connection between the predicate and the

concept of the subject. To be sure, in the judgment the predicate ( $c$ ) is connected with the subject concept ( $a + b$ ); but the connection is grounded in, and mediated by, the reference of both to the identical object ( $x$ ). Consequently, it extends our knowledge of  $x$  (in this case, of all  $x$ 's) by providing a determination or property of  $x$  that is not already contained in the concept ( $a + b$ ). This is what is meant by a "material extension."

Kant explicates this further by suggesting that the synthetic judgment contains a "determination," whereas the analytic judgment contains only a "logical predicate" (JL 9: III; 607). Since Kant maintains both that existential judgments are synthetic and that 'existence' is not a real predicate, this account of synthetic judgments cannot be accepted as it stands. In other words, it cannot be maintained that the possession of a logical predicate that is also a real predicate is criterial for the syntheticity of a judgment. An existential judgment is synthetic not because its logical predicate 'existence' is a real predicate or determination but rather because its logical subject is one, and the judgment simply asserts the existence of an object corresponding to this subject.

It might also seem that in analytic judgments, such as "All bodies are divisible," the logical predicate, 'divisibility', is likewise a real predicate, because it is a property of every  $x$  that answers to the general description thought in the concept 'body'. Nevertheless, the point is that in analytic judgments the predicate is related to the subject ( $x$ ) simply in virtue of the fact that it is already contained (either immediately or mediately) in the concept of this subject. Thus, the "reality" of the predicate does not come into consideration. In synthetic judgments, however, the reference to the subject and, therefore, the reality of the predicate are just the points at issue. That is why the question of how such judgments are possible *a priori* is so perplexing. 

In any case, a synthetic judgment (of theoretical reason)<sup>39</sup> can materially extend our knowledge only if the concepts in it are related to intuition. The reason for this is to be found in the very nature of discursive thought. As we have already seen, concepts can never relate immediately to objects, only to other representations (either concepts or intuitions). Consequently, no judgment can ever relate a concept directly to an object, only to some given representation of it. But if the concept is held to be a real predicate or determination, then it must be related to some representation that itself stands in an immediate relation to the object, that is, to an intuition. In fact, it is only if the subject and predicate concepts in a synthetic judgment are *both* related to the intuition of the object that the connection of these concepts thought in the judgment can be grounded or objectively valid.<sup>40</sup>



#### IV. THE PROBLEM OF THE SYNTHETIC *A PRIORI*

Kant held that the question of the possibility of synthetic *a priori* judgments emerges as the central problem of metaphysics as soon as the distinction between analytic and synthetic judgments is properly drawn. He thus points to the failure of past philosophers to recognize this problem as evidence of their failure to make the distinction. No such claim, however, is made for the *a priori*–*a posteriori* distinction. Indeed, in his response to Eberhard Kant admits that this distinction is “long known and named in logic” (UE 8: 228; 318). To ask whether a given judgment or proposition is *a priori* or *a posteriori* is to ask how it is known to be true or, in Kantian terms, how it is grounded or legitimated. In short, much like ‘*an sich selbst*’ discussed in the preceding chapter, these terms function adverbially to characterize *how* something is known rather than substantively to characterize *what* is known (for example, *a priori* truths).

The key issue is the role of experience in this grounding. *A priori* judgments are those that are grounded independently of experience, while *a posteriori* ones are grounded by means of an appeal to it. Following Leibniz, Kant regards necessity and universality as the criteria for the *a priori*. Thus, he shares with his predecessor, as well as many other philosophers, the assumption that the truth value of judgments laying claim to universality and necessity cannot be grounded empirically.

Since their truth value is determined by means of an analysis of the constituent marks of a given concept, analytic judgments clearly fit into this category. Moreover, this is true even when the concept is empirical. The real question is whether it is likewise possible for synthetic judgments to have non-empirical grounds. On the one hand, as synthetic, they cannot have a purely conceptual or logical grounding; while, on the other hand, as *a priori*, they cannot be grounded in experience. Seen from a Kantian perspective, the problem of the synthetic *a priori* is, therefore, that of explaining how a non-empirical, yet extra-conceptual and extra-logical, grounding of a judgment is possible. An equivalent way of formulating the problem is to ask how it is possible to extend one’s knowledge (in the material sense) beyond a given concept, independently of any experience of the object thought under that concept.

Perhaps Kant’s clearest answer to this question is in a text not published in his lifetime. As he there puts it:

Knowledge is a judgment from which proceeds a concept that has objective reality, i.e., to which a corresponding object can be given in experience. But all experience consists of the intuition of an object, i.e., an immediate and individual representa-

tion, through which the object is given as to knowledge, and of a concept, i.e., a mediate representation through a characteristic common to several objects, whereby it is therefore thought. Neither of these two modes of representation constitutes knowledge on its own, and if there are to be synthetic *a priori* cognitions, there must also be *a priori* intuitions as well as concepts. [Fort 20: 266; 358–59]

The essential point here is the claim that synthetic *a priori* judgments require *a priori* or pure intuitions as well as concepts. The need for the latter is easily shown. If the predicate in a synthetic judgment were an empirical concept, its connection with the object would have to be established empirically, which would make the judgment empirical. As obvious as this point is, however, it is somewhat obscured by two features of Kant's position: one is that analytic judgments, which are always *a priori*, can be made on the basis of empirical concepts; the other is that Kant sometimes speaks of "impure" *a priori* judgments, which involve empirical concepts. With regard to analytic judgments, we need only note once again that they abstract from the whole question of objective reference and thus from the objective reality of the concept. Consequently, the appeal to experience is otiose, even when the concept itself is empirical. With regard to "impure" *a priori* judgments, it is sufficient to note that they always involve pure concepts as predicates. In Kant's own example, "Every alteration has a cause" (B3), the emphasis is placed on the fact that 'alteration' is an empirical concept. But 'causality' is a pure concept, which is what makes the judgment *a priori*.

The role of pure intuition in synthetic *a priori* judgments, which is the point on which Kant particularly insists in all of his anti-Leibnizian moments, is considerably more complex. It involves at least three questions, each of which must be considered separately. The first and most basic is why synthetic *a priori* judgments require intuitions at all. The second is why they require *pure* rather than merely empirical intuitions. The third is whether such judgments can be said to relate pure concepts to pure intuitions or, equivalently, whether they require that pure intuitions be "subsumed" under pure concepts.

Although the answer to the first question is obvious, it is worth emphasizing. A synthetic *a priori* judgment requires intuition for the same reason that any synthetic judgment does: the mutual reference to intuition of the concepts connected in a judgment is what alone makes possible the material extension of knowledge. Indeed, it is precisely because of the impossibility of providing an intuition answering to the concepts that Kant holds the judgments of transcendent metaphysics to be ungrounded. From the standpoint of theoretical knowl-

edge at least, the limits of our sensibility (the source of all of our intuitions) are at the same time the limits of our world.

The question thus becomes why we cannot make do with empirical intuition. Why is it necessary to introduce the deeply perplexing notion of a pure, yet sensible, intuition, which Eberhard and so many others have dismissed as a contradiction in terms? Although the notion of a pure intuition is a murky one, the need for it is clear: the insufficiency of empirical intuition to ground a synthetic *a priori* judgment. The problem with empirical intuition lies in its particularity. A determinate empirical intuition is the representation of a particular empirical object: for example, the desk in front of me. As particular, the representation is incapable of expressing the universality and necessity that is thought in a pure concept and asserted in a synthetic *a priori* judgment. To cite a mathematical example: as synthetic, the judgment that the sum of the three interior angles of a triangle is equal to two right angles must be grounded in the intuition of a triangle; but as *a priori*, it cannot be grounded in the intuition (image) of any particular triangle. Its possibility thus rests upon there being some non-empirical or pure intuition of “triangularity as such,” that is, a singular representation that nonetheless can “attain the generality of the concept, which makes this valid for all triangles, right or acute, etc.” (A141 / B180).<sup>41</sup>

Finally, we come to the question of whether pure intuitions function as conditions of synthetic *a priori* judgments in the same way that empirical intuitions function in those that are cognized *a posteriori*. At this stage of the analysis, however, before we have investigated Kant’s conception of pure intuition, it can only be argued in general terms that the very same considerations advanced in support of the contention that synthetic judgments in general require the relation of concepts to intuitions (or, equivalently, the “subsumption” of the latter under the former) must also apply to the connection between those that are cognized *a priori* and pure intuition. How, after all, could a pure concept apply universally and necessarily to a sphere of objects, as it must if the judgment is to be both synthetic and *a priori*, unless it were related in the judgment to the universal and necessary conditions, that is, the “form,” of our intuition of these objects? But these universal and necessary conditions of intuition are, as we shall see, themselves pure intuitions. Consequently, if synthetic judgments are to be possible *a priori*, pure concepts as (real) predicates must be related in these judgments to pure intuitions.

## Chapter 5 The Sensible

### Conditions of Human Cognition

At the beginning of the Transcendental Aesthetic, after defining some key terms and linking space with outer sense, defined as “a property of our mind” through which “we represent to ourselves objects as outside us, and all as in space,” and time with inner sense, defined as the means by which “the mind intuits itself, or its inner state” (A22 / B37), Kant turns abruptly to the question of the nature of space and time. Four possibilities are introduced. They might be: (a) actual entities (substances), (b) determinations of things (accidents), (c) relations of things that “would pertain to them even if they were not intuited,” or (d) “relations that only attach to the form of intuition alone, and thus to the subjective constitution of our mind, without which these predicates could not be attached to anything at all” (A23 / B37–38).<sup>1</sup>

Since the first three constitute the traditional ontological options and the fourth Kant’s critical view, it is clear that these alternatives are meant to be exhaustive. Nevertheless, the situation is complicated by the fact that Kant had already indicated in the *Inaugural Dissertation* that he considered there to be only two serious alternatives to his own view: the Newtonian thesis that space is “an *absolute* and boundless *re-*

*ceptacle* of possible things” and the Leibnizian contention that it is the “relation itself which obtains between existing things, and which vanishes entirely when the things are taken away, and which can only be thought as being between actual things” (Diss 2: 403–4; 397).<sup>2</sup> Although this appears to ignore it, option (b), actually comes closer than (a) to characterizing the Newtonian view. For in spite of affirming the absolute reality of space and time, Newton denies (largely on theological grounds) that they are substances, asserting instead that they are accidents of God.<sup>3</sup> Thus, it is the view that space and time are substances that drops out of the picture.<sup>4</sup> Or, more precisely, Kant effectively elides (a) and (b), leaving only the Newtonian and Leibnizian views as serious alternatives.

Be that as it may, the demonstration of the correctness of the fourth alternative, which Kant presents as equivalent to a direct proof of transcendental idealism, is the main goal of the Transcendental Aesthetic. And since the alternatives to be rejected are all ontological in nature, it would seem that the Kantian alternative must likewise be so, that is, a thesis about the *real* nature of space and time. Moreover, if this is the case, then it appears that transcendental idealism must itself be regarded as an ontological thesis, in direct competition with its transcendentially realistic alternatives.

Although this is the way in which Kant’s argument is usually construed—indeed, how it must be construed if viewed through transcendentially realistic spectacles—the results of the first three chapters should make one wary of drawing such a conclusion. Uncontroversially, Kant’s doctrine of transcendental ideality involves a denial of the traditional ontologies of space and time (the alternatives available to transcendental realism), but it does not follow from this that it is itself an alternative ontology. It may also be seen as an *alternative to ontology*, according to which space and time are understood in terms of their epistemic functions (as forms or conditions of outer and inner sense, respectively) rather than as “realities” of one sort or another.

The twin aims of this chapter are to substantiate such a reading and to defend the basic thrust of Kant’s argument, so interpreted. The chapter is divided into two main parts. The first deals with Kant’s analyses of the *a priori* and intuitive nature of the representations of space and time and their function as sources of *a priori* knowledge. The second interprets and evaluates the conclusions Kant draws from these analyses. It is there that we shall consider the issue of transcendental idealism and Kant’s argument for it.

## I. THE REPRESENTATIONS OF SPACE AND TIME

In the second edition, Kant divides his analysis of the representations of space and time into what he terms “metaphysical” and “transcendental expositions.” By an exposition he understands “the distinct (even if not complete) representation of what belongs to a concept.”<sup>5</sup> It is metaphysical “when it contains that which exhibits the concept as given *a priori*” (B38); it is transcendental if it provides “the explanation of a concept as a principle from which insight into the possibility of other synthetic *a priori* cognition can be gained” (B40). Simply put, the task of a metaphysical exposition is to establish the *a priori* origin of a certain representation, while that of a transcendental exposition is to show how this representation grounds the possibility of other synthetic *a priori* knowledge.

The metaphysical expositions have two goals, both of which concern the question of origin. The first is to show that the representations of space and time are *a priori*, that is, that they have a non-empirical origin. The second is to show that this origin lies in sensibility, because they are intuitive rather than conceptual in nature. The discussion as a whole, however, is divided into four parts. The first two deal, respectively, with the apriority and intuition theses of the metaphysical expositions, the third with the thorny problem of givenness, that is, how space and time may said to be “given” in *a priori* intuition independently of the conceptual activity of the understanding, and the fourth with the transcendental expositions and their implications for idealism. Since the expositions of space and time parallel one another, we shall concentrate on the former, referring to the latter only in those instances where Kant’s analysis differs significantly.

### A. The Apriority Thesis

The argument for the *a priori* nature of the representation of space is contained in two brief paragraphs, which have become the topic of endless controversy.

Space is not an empirical concept that has been drawn from outer experiences. For in order for certain sensations to be related to something outside me (i.e., to something in another place in space from that in which I find myself), thus in order for me to represent them as outside and next to one another, thus not merely as different but as in different places, the representation of space must already be their ground. Thus the representation of space cannot be obtained from the relations of outer appearance through experience, but this outer experience is itself first possible only through this representation. [A23 / B38]

Space is a necessary representation, *a priori*, that is the ground of all outer intuitions. One can never represent that there is no space, although one can very well think that there are no objects to be encountered in it. It is therefore to be regarded as the condition of the possibility of appearances, not as a determination dependent on them, and is an *a priori* representation that necessarily grounds outer appearances. [A24 / B38–39]

The problems begin with the question of the relationship between the two arguments. Are they two independent proofs of the same thesis; one argument with two steps; or two distinct arguments concerned with two distinct conceptions of *apriority*? Since Kant is silent on the topic, it is not surprising that the commentators are divided.<sup>6</sup> The position taken here is that they constitute two independent proofs, operating with the same conception of *apriority*. Nevertheless, they embody two quite distinct proof strategies, and the second, though not making a significantly stronger claim, calls attention to a crucial feature of the representation of space that is ignored by the first.<sup>7</sup>

*The First Apriority Argument.* This argument involves two distinct claims, both of which are presuppositional in nature. The first is that the representation of space must be presupposed, if I am to refer my sensations to something “outside me” (*ausser mir*). The second is that it must be presupposed if I am to represent objects as outside or external to one another. Accordingly, each claim refers to an epistemological task that the representation of space performs and that it presumably could not perform if it were “drawn from outer experiences” or, as Kant puts in the conclusion, “obtained [through experience] from the relations of outer appearance.”

The first question that arises concerns the meaning of ‘*ausser*’. Since this is usually understood in a spatial sense, the claim that space must be presupposed in order to refer my representations (sensations) to something *ausser mir* might appear to be tautologous. Moreover, a similar objection might be raised against the second claim as well, since to assert that objects are outside one another is just to say that they occupy different regions of space. If we focus on Kant’s initial characterization of outer sense, however, the claims appear in a somewhat different light. For, as I have already noted, Kant asserts that through outer sense “we represent to ourselves objects as outside us *and* [my emphasis] all as in space.” Here ‘outside us’ clearly means external to or distinct from ourselves and our states; while the fact that such objects are *also* represented as in space is viewed as a specific feature of our outer experience. Similarly, by ‘inner sense’ is meant a sense through which one becomes perceptually aware of the self and its

inner states. Consequently, Kant's claim that the representation of space must be presupposed in order to be aware of things as *ausser uns* (or each other) is no more tautological than the corresponding claim about time.

Nevertheless, it would be a mistake to take the first argument as turning on this non-spatial sense of '*ausser*.'<sup>8</sup> Indeed, Kant makes it quite clear that by the relation to something *ausser mir* he means not simply something distinct from the self and its states but "something in another place in space from that in which I find myself." Moreover, the second part of the argument claims not that the representation of space is required to distinguish objects from one another (purely qualitative features could do that),<sup>9</sup> but rather that it is necessary in order to represent them "as not only different but as in different places." Thus, the argument involves a peculiar mixture of the spatial and non-spatial senses of '*ausser*,' which turns out to be a source of considerable confusion.

As a first step in clarifying the situation, it is crucial to recognize that the confusion (if there be any here) is not Kant's. For the initial characterization of outer sense calls attention to this very distinction, and the subsequent clarification of the meanings of '*ausser*' does so as well. This suggests that it is the empiricist who is guilty of confusion in attempting to derive the representation of space from a prior awareness of things as outside me (my body) and/or as external to one another. In fact, although Kant does not make the point explicitly, there is reason to believe that he may be at least hinting at the idea that it is the empiricist's conflation of the two senses of '*ausser*' that explains, at least in part, the failure to recognize the circularity of such attempts.<sup>10</sup>

This circularity becomes evident when the empiricist account is framed consistently in spatial terms. For it is then apparent that the representation of space cannot be derived either from a prior awareness of things as outside me (my body) or from the perception of things as external to one another, because the thought of either relation presupposes the representation of space. In the case of time, the parallel claim is that the awareness of the relations of simultaneity and succession presuppose the representation of time (A30 / B46). Accordingly, the attempt to account for the origin of our representations of space and time in these ways may be dismissed as inherently question begging. In endeavoring to describe the experience through which the mind acquires these representations, the empiricist tacitly assumes that the mind already has them.

Kant's argument is neither trivial nor lacking in epistemic force. It is not trivial because the claim is not simply that space must be presupposed in order to represent things as spatial, but rather that it must be presupposed as a condition of the possibility of the perception of the relations from which the empiricist



account claims it is derived. It is not lacking in epistemic force because it indicates that the representation of space is such a condition precisely by functioning as a “ground” of the representation of these relations. In fact, Kant himself underscores the epistemic thrust of the argument by concluding that “this [*diese*] outer experience is itself first possible only through this representation” (A23 / B38). In other words, for Kant this argument of itself shows not simply that the representation of space is *a priori* but that it is such precisely because it serves as a condition of outer experience.<sup>11</sup>

Although the focus of this argument is anti-empiricistic, it also applies to the relational theory of Leibniz, particularly as it is articulated in the correspondence with Clarke. For in spite of his rationalism, Leibniz’s views on the epistemology of space and time are basically empiricistic. Thus, after contending against the Newtonian theory that space is simply the order of coexisting phenomena, he abruptly turns in his fifth letter to the question of “how men come to form to themselves the notion of space.” According to Leibniz, the mind first notices a certain order (situation or distance) among coexisting things, which is equated with the set of relations in which these things stand to one another. This, he suggests, is sufficient to account for the idea of place (actually, the same place), which, in turn, leads to the idea of space, understood as “that which comprehends all those places.” And he takes this as showing that “in order to have an idea of place, and consequently of space, it is sufficient to consider these relations and the rules of their changes, without needing to fancy any absolute reality out of the things whose situation we consider.”<sup>12</sup>

Kant agrees with Leibniz that there is no need to attribute absolute reality to space, as well as with Leibniz’s subsequent contention that space itself is “merely an ideal thing.”<sup>13</sup> Nevertheless, this should not obfuscate the deep disagreement between them on this issue. First, as should be apparent from my earlier discussions, Kant understands this ideality in a way far different from Leibniz. Second, and of more immediate concern, the first apriority argument contains a direct denial of the Leibnizian epistemology of space. The main thrust of this argument is that the representation of order or relation, which for Leibniz is the primary datum for constructing the notions of place and space, must already be understood as an order or relation of things *in space*. As a result, any such order (of situation or distance) presupposes space as the condition of its possibility. Thus, Kant reverses the Leibnizian account of epistemic precedence. Whereas Leibniz gave primacy to the perceived order or relation of things, with this being sufficient to define a thing’s place, and the general no-

tion of space being understood as the network of places; for Kant it is the representation of space that is primary.

These considerations put us in a position to deal with two fairly common lines of objection to Kant's argument. The first can be traced back to Eberhard's collaborator, J. G. Maass, and it reflects the latter's endeavor to defend the Leibnizian position against Kant's attack. According to Maass, it is possible to accept Kant's premises and deny his conclusion. A representation *A*, he reasons, may underlie or be presupposed by another representation, *B*, and thus not be derivable from it. But it does not follow from this that *A* is *a priori*. An equally plausible alternative is that the representations are correlative and that they mutually condition one another. Assuming this possibility, which Maass thinks Kant ignores, the concept *A* could be obtained only by abstraction from the complete concept *AB*, which would make it an empirical concept.<sup>14</sup> This is not merely a hypothetical alternative but a statement of the Leibnizian position. Thus, *A* refers to the order or situation of things and *B* to the things themselves. The point, then, is that, while we cannot represent the things (*B*) without also representing their order or situation (*A*), we only arrive at the concept of the latter by an act of abstraction from the complete concept (*AB*), which makes it empirical.<sup>15</sup>

Insofar as it assumes with Leibniz that space is to be understood as an order or relation of things and argues from this that our representations of this order and of the things ordered mutually condition one another, this objection begs the question against Kant. For Kant's claim is that the representation of this order or relation presupposes the representation of space and, therefore, that the former cannot be derived from the latter.<sup>16</sup> Nor can this be evaded by maintaining that the reciprocal connection is between the representation of the order of things and the representation of space, because the whole thrust of the argument is that the latter is the condition of the former and not vice versa.

Whereas the first line of objection maintains that Kant's argument does not prove enough, the second holds that it proves too much. According to this objection, if the argument proves anything at all, it is that even our empirical concepts must be *a priori*. This rests on the assumption that Kant's argument moves directly from the claim that the representation of space is necessary in order to recognize spatial relations and determinations to the conclusion that it is *a priori*. It is then suggested that an analogous claim can be made about any empirical concept. D. P. Dryer, who himself rejects this objection, describes it thus:

In order to recognize objects as red, one must already have the concept of redness. But this does not establish that the concept of red is not an empirical concept. In order to observe things about us, we must represent them in space. How then can this show that the concept of space is not likewise empirical?<sup>17</sup>

Although its thrust is quite different, this objection fails for essentially the same reason as the first: it misconstrues Kant's claim about the connection between the representation of spatial relations and the representation of space. Thus, it takes Kant to be saying merely that the representation of objects as spatially related presupposes the representation of spatial relations. And given this assumption, it concludes that a parallel claim can be made about the connection between any concept and the items falling under it. In reality, however, Kant's argument turns on the *disanalogy* between the two cases. For on his view, spatial relations are not instances of the concept of space from which the latter is formed by abstraction or construction. It is rather that these relations must be conceived as *in* and with reference to space, which cannot be said of the relation between a concept and the instances falling under it.<sup>18</sup>

*The Second Apriority Argument.* This argument asserts that "space is a necessary representation, *a priori*, that is the ground of all outer intuitions." It is based on two premises contained in the second sentence: "One can never represent that there is no space [*Man kann sich niemals eine Vorstellung davon machen dass kein Raum sei*], though one can very well think that there are no objects to be encountered in it." This suggests the following argument schema: If *x* can be (or be represented) without *A*, *B*, *C* and their mutual relations, but *A*, *B*, *C* cannot be (or be represented) without *x*, then *x* must be viewed as a condition of the possibility of *A*, *B*, *C* and their mutual relations (or the representation thereof).<sup>19</sup> Applying this to space, Kant concludes, "It must therefore be regarded as the condition of the possibility of appearances, not as a determination dependent upon them, and is an *a priori* representation that necessarily grounds outer appearances."

The first premise is frequently dismissed on the grounds that it involves a psychological claim (and a questionable one at that). Thus, according to Kemp Smith, "the criterion is not the impossibility of thinking otherwise, but our own incapacity to represent this specific element as absent."<sup>20</sup> Other commentators, however, suggest that Kant is asserting a logical rather than a psychological impossibility. On this interpretation, Kant is making a claim about the in-


conceivability of the nonexistence of space that is analogous to Spinoza's claim about the inconceivability of the nonexistence of substance.<sup>21</sup>

Although the psychological reading is the most natural, it dooms the argument from the beginning. For if it is interpreted as asserting the impossibility, as a matter of psychological fact, of representing to oneself the absence of space, it becomes difficult to see how it could support, much less entail, Kant's conclusion that "space is an *a priori* representation that *necessarily grounds* [my emphasis] outer appearances." Granted, this does not rule out such a reading, since we cannot assume that Kant is incapable of such a *non sequitur* (particularly since he is frequently accused of similar ones); but it does suggest that it should be accepted only as a last resort.

The logical reading is also unacceptable, however, on both textual and doctrinal grounds. First, Kant nowhere affirms that space (or time) is logically necessary. On the contrary, we have seen that it is not necessary for space to be the form of outer or time of inner sense, since Kant holds open conceptual space for beings with other forms of sensibility. Second, Kant's doctrine that we can think, though not know, things as they are in themselves requires him to allow for this possibility. For how could we even think about things as they are in themselves unless we could "represent that there is no space"?

Fortunately, there is a third alternative, which accords with both the wording of the text and Kant's overall views, while also making it possible to see how the premise contributes to the establishment of the desired conclusion. According to this reading, the claim is that we cannot represent outer appearances without also representing them as in space.<sup>22</sup> In other words, the ineliminability of space is applied to the outer intuition of beings with our sensory forms and in no way precludes the possibility either of other beings, with different forms of outer intuition, or of our thinking (though not intuiting) the absence of space with respect to things as they are in themselves. Although this claim may be described as broadly psychological, since it concerns the cognitive capacities of beings like ourselves, it is also epistemological because it asserts that space is necessary for the representation of outer appearances.

Additional textual support for this reading is found in the parallel argument regarding time, where Kant writes: "In regard to appearances in general one cannot remove [*aufheben*] time, though one can quite well take the appearances away from time" (A31 / B46). He further remarks, "In it [time] alone is all actuality of appearances possible." Finally, in a parenthetical phrase added in the second edition, time is characterized as "the universal condition of their

[appearances] possibility” (B46). The significance of this lies in the fact that the scope of the claim concerning “removing time” (the temporal analogue to representing the absence of space) is limited to appearances. But we are also told in the conclusion that the reason we cannot remove time from appearances is that time is a condition of their very possibility, that is, their representation. This is an epistemological, not a psychological, claim. 

Making allowance for the scope distinction between the two claims, which is a consequence of the fact that all appearances are in time, whereas only outer appearances are in space, it seems reasonable to assume that Kant is trying to make the same point about the representation of space. So construed, the claim is not that we cannot *in any sense* “remove space” but merely that we cannot do so without also removing all outer appearances. Kant put the matter best in the observation on the antithesis of the First Antinomy, where, in commenting on the Leibnizian project of reconciling the finitude of the world with the rejection of a void, extra-mundane space, he remarks, “If one wants to . . . leave out space in general as the *a priori* condition of the possibility of appearances, then the whole world of sense is left out” (A433 / B461).

Nevertheless, it does not follow from the fact that we cannot think of appearances without also thinking of them as in space and time, that the latter are *a priori*. It might also be the case that we cannot think about space and time without also thinking about appearances. In that case, however, these representations would be empirical, formed by abstraction from the complex representation of appearances as situated in space and time. Indeed, this is where Maass’s objection, mistakenly presented as a criticism of the first argument, becomes relevant. But the conclusion to be drawn is simply that the second part of the claim is equally necessary for the establishment of the apriority thesis. In other words, it is necessary to show *both* that we cannot remove space and time from the thought of appearances *and* that we can remove appearances from the thoughts of space and time. For these premises jointly entail that the representations of space and time are conditions of appearances and thus *a priori*.

Clearly, then, the task is to understand the latter claim, which appears to involve difficulties comparable to those encountered in the first part of the argument. To begin with, it cannot be taken to mean that we can somehow perceive or experience empty space or time, since Kant frequently denies this to be possible.<sup>23</sup> Moreover, even if it were possible, it is hard to see what bearing it could have on the apriority issue.

A more viable alternative is that Kant is conducting a thought experiment in which one removes the appearances from space and time, just as in the first step

in the argument one attempted (unsuccessfully) to remove space and time from the appearances. That this is Kant's intent is strongly suggested by his statement in the time version of the argument that "one can very well take the appearances away from time" (A31 / B46). Moreover, Kant reports the result of such an experiment in the *Prolegomena*, when he remarks, "[I]f one eliminates from the empirical intuition of bodies and their alterations (motion) everything empirical, that is, that which belongs to sensation, space and time still remain" (Pro 283; 79).

In the *Critique*, he expresses himself more expansively on the matter with respect to space:

So if I separate from the representation of a body that which the understanding thinks about it, such as substance, force, divisibility, etc., as well as that which belongs to sensation, such as impenetrability, hardness, color, etc., something from this empirical intuition is still left for me, namely extension and form. These belong to pure intuition, which occurs *a priori*, even without an actual object of the senses or sensation, as a mere form of sensibility in the mind. [A20–21 / B35]

This much discussed passage from the introductory portion of the *Aesthetic* is best seen as a statement of what Kant hopes to show rather than as a self-contained argument. Accordingly, like the comparable passage from the *Prolegomena*, it gestures toward the thesis that the representation of space is a pure intuition. Nevertheless, it also serves to illustrate the more basic point that if one abstracts from both the conceptual and sensory content of the empirical representation of a body, the extension and form or figure [*Gestalt*] of this body, that is, its purely spatial features, remain. These remain not in the sense that they can be perceived independently of the sensory content but that they provide a representational content with a determinate structure (presumably topological, affine, and mereological),<sup>24</sup> which is not dependent on this sensory content.

What is of immediate relevance, however, is the asymmetry between the sensory and the spatial content of our empirical representation of 'body'. As the thought experiment indicates, although the latter remains when one abstracts from the former, the converse does not hold. Thus, rather than their being equal partners, the latter is epistemically prior in the sense that it is the *a priori* condition of the representation of the former.

Although this seems to be a stronger claim than the one made in the first argument, if one keeps in mind the first's conclusion that "this outer experience is itself first possible only through this representation," the difference becomes much less significant. The question is only whether the above expression refers

to outer experience in general (in which case the two claims are precisely equivalent) or merely to the experience of the relations of outer appearances with which the first argument is concerned (in which case the present argument involves a somewhat stronger claim). But if, as seems reasonable, one also assumes that these relations are themselves conditions of the representation of outer appearances, then these conclusions amount again to much the same thing: the representation of space is *a priori* in the sense of being a condition of outer appearances rather than being either derivable from or coequal with the latter.<sup>25</sup>

Nevertheless, this does not make the second argument redundant, since it calls attention to a feature of these representations that is both central to Kant's position and completely neglected by the first argument: the fact that they have a content of their own, which remains when abstraction is made from everything empirical. By bringing this out Kant not only underscores their *a priori* status but also helps prepare the way for the claim that they are pure intuitions. It is to this claim that we now turn.

## B. The Intuition Thesis

Once again Kant offers two arguments in support of a single thesis. This time, however, the situation is complicated by the fact that he substitutes a completely different version of the second argument in the second edition. For convenience sake, both versions of the second argument will be cited, but the analysis will focus on that in the second edition.

Space is not a discursive or, as is said, general concept of relations of things in general, but a pure intuition. For, first, one can only represent a single space, and if one speaks of many spaces, one understands by that only parts of one and the same unique space. And these parts cannot as it were precede the single all-encompassing space as its components (from which its composition would be possible), but rather are only thought in it. It is essentially single; the manifold in it, thus also the general concept of spaces in general, rests merely on limitations. From this it follows that in respect to it an *a priori* intuition (which is not empirical) grounds all concepts of them. Thus also all geometrical principles, e.g., that in a triangle two sides together are always greater than the third, are never derived from general concepts of line and triangle, but rather are derived from intuition and indeed derived *a priori* with apodictic certainty. [A24–25 / B39]

Space is represented as a given infinite magnitude. A general concept of space (which is common to a foot as well as an ell) can determine nothing in respect to magnitude. If there were not boundlessness [*Grenzenlosigkeit*] in the progress of in-

tuition, no concept of relations could bring with it a principle of their infinitude. [A25]

Space is represented as an infinite given magnitude. Now one must, to be sure, think of every concept as a representation that is contained in an infinite set [*Menge*] of different possible representations (as their common mark), which thus contains these under itself; but no concept, as such, can be thought as if it contained an infinite set of representations within itself. Nevertheless space is so thought; (for all the parts of space, even to infinity, are simultaneous). Therefore the original representation of space is an *a priori* intuition, not a concept. [B39–40]

*The First Intuition Argument.* This argument assumes the exhaustiveness of the concept-intuition distinction. Given this assumption, it attempts to prove, by means of an analysis of the nature of the representation of space, that as it cannot be a concept it must be an intuition. Or, more precisely, since we shall see that Kant also acknowledges concepts of space (and time), it endeavors to show that “the original representation of space” (B40) is intuitive rather than conceptual in nature.<sup>26</sup> Having already established that the representation is “pure” through the *apriority* arguments, he now concludes that it is a pure intuition and spells out the implications of this result for geometry (a point that may be ignored for the present). The actual proof consists of two steps. In the first, Kant contrasts the relation between space and particular spaces with that between a concept and its extension, and in the second with the relation between a concept and its intension.<sup>27</sup> Once again, both steps are required to produce the desired conclusion.

The first step turns on the singleness of space. The basic claim is that “one can represent only a single space.” If the argument is to work it must be assumed that this is not a contingent matter, as if the class of spaces just happened to have only one member. But neither can it be a logically necessary truth, like the truth that we can conceive of only one “most perfect being.” In support of this claim, however, Kant offers only the observation that we are somehow constrained to think of particular spaces as parts of a single space. Although Kant asserts in the parallel time argument that a “representation . . . which can only be given through a single object, is an intuition” (A32 / B47), this still does not prove that the representation of space (or time) is an intuition. Consider, for example, the concept of the world, which underlies the argument of the Antinomies. Since it is the concept of a complete collection or totality, we can conceive of only one (actual) world.<sup>28</sup> Nevertheless, it does not follow from this that the representation is an intuition.<sup>29</sup> Thus, in order to prove that the original repre-



sensation of space is intuitive, Kant must show how it differs from the concept of a complete collection or totality, such as that of the world.

Although it is unclear whether Kant actually had this problem in mind, it is effectively resolved in the second part of the argument, where he contrasts the relationship between space and its parts (particular spaces) with the relationship between a concept and its intension. The main point is that the marks or partial concepts out of which a general concept is composed are logically prior to the whole. But this is not the case with space and its parts. Rather than being pre-given elements out of which the mind forms the idea of a single space, the parts of space are only given in and through this single space that they presuppose. Space, in other words, is presented not only as single (*einzig*) but also as a unity (*einig*). Consequently, it cannot be conceived of as a collection or composite of spaces, any more than it can be viewed as an ordinary class concept.

As Kant proceeds to make clear, however, this result does not conflict with the fact that we can form general concepts of space. Kant is perfectly happy to admit this, but he insists that such concepts arise from the limitation of the one, all-inclusive, and immediately given space.<sup>30</sup> Here he seems to have had in mind a two-step procedure: first, through the introduction of limitations (itself a conceptual activity) we produce the idea of determinate spaces (figures and magnitudes); then, on the basis of these determinations, we form by abstraction general concepts of spaces. But the main point is simply that the possibility of such conceptualization rests upon some given content, that is, an intuition. In Kant's own terms, space is *intuitus, quem sequitur conceptus*.<sup>31</sup>

*The Second Intuition Argument.* This argument is both more complex and more problematic than the first. It assumes that space is represented as an infinite *given* magnitude and concludes from this that the representation must be an intuition. Both the reference to infinity and the emphasis on the givenness of space differentiate this argument from the previous one and bring it closer to the Kantian conception of intuition. Moreover, the latter point highlights an essential feature of Kantian intuitions that is largely ignored in the first. But it also raises problems about givenness, which will be taken up shortly.

Leaving that aside for the present, a brief glance at the first-edition version of the argument makes it clear why Kant completely recast it in the second. Its nerve is the claim that a general concept of space, which is formed by abstraction from particular spatial measurements, "could not determine anything with regard to magnitude." This is obviously true, but it is irrelevant, since it does not have anything in particular to do with the *infinity* of space. Precisely


the same point could be made if space were represented as a finite given magnitude.

The second-edition version avoids this difficulty by offering an analysis of the different senses in which concepts and intuitions involve infinity. Moreover, in so doing, it sheds additional light on the differences between the form or structure of concepts and intuitions.<sup>32</sup> A concept has a complex structure, involving both an extension and an intension. Viewed extensionally, every concept has various other concepts falling *under* it. These are arranged hierarchically in terms of generality, and they stand in the relationship of genus to species. More specific concepts are introduced by adding differentia. Thus, the genus “physical body” can be divided into the species “inanimate” and “animate,” and the latter into “animal” and “vegetable,” and so on. Viewed intensionally, every concept contains other concepts *within* it as its component parts. But here the ordering is the reverse of the extensional ordering, since the lower or more specific concepts contain the higher or more general ones within themselves. In short, there is an inverse correlation between the extension and the intension of a concept. The smaller the extension, the richer the intension, and vice versa.





As the first intuition argument makes clear, this contrasts markedly with the structure of an intuition. Since it is the representation of an individual, all of its components are contained in and presuppose the whole. Similarly, intuitions are divided not by adding differentia but by introducing limitations or boundaries.<sup>33</sup> The second argument builds on this by showing how the difference in structure is reflected in the different ways in which concepts and intuitions involve infinity. A concept does so extensionally: it can have an infinite or, better, an indefinite number of concepts falling under it. In fact, since Kant denies that there can be an *infima* species, he is committed to the view that the search for subordinate concepts can be pursued ad infinitum. A concept cannot have an infinitely rich intension, however, because such a “concept” would lose its discursive character. Although an intuition does not have an intension, it could have an infinite number of parts within it.<sup>34</sup> And, Kant suggests, this is precisely the way in which space is thought, “for all the parts of space, even to infinity [*ins Unendliche*], are simultaneous.” From this he concludes that “the original representation of space is an *a priori* intuition, not a concept.”

One fundamental question raised by this analysis concerns the sense of ‘infinity’ assigned to space. Although it is hardly a model of clarity, the claim that the parts of space, “even to infinity, are simultaneous” suggests that Kant has in mind the innumerability of its parts.<sup>35</sup> But this conflicts both with his denial


that space is a composite and with the thesis of the First Antinomy, which rules out the possibility of attributing infinity in this sense to the spatiotemporal world.<sup>36</sup> Consequently, if Kant's claim is to be made consistent with what he says elsewhere, some other sense of the term must be operative in the present argument. 

Fortunately, there is considerable evidence to suggest that this is the case. First, in the first-edition version of the argument, Kant makes it clear that the infinity of space has to do with the "boundlessness in the progress of intuition." In other words, however large a region of space one takes, it is always represented as bounded by more of the same (presumably, the same applies, *mutatis mutandis*, to its division).<sup>37</sup> Second, in the parallel time argument Kant asserts that "the infinitude of time signifies nothing more than that every determinate magnitude of time is possible only through limitations of a single time grounding it. The original representation time must therefore be given as unlimited [*uneingeschränkt*]" (A32 / B47–48). Since there is nothing to indicate that Kant construed the infinity of space any differently from that of time, it seems reasonable to assume that he took both to consist in their unboundedness or limitlessness.<sup>38</sup>

### C. The Givenness of Space (Form of Intuition and Formal Intuition)

Quite apart from worries about the sense of infinity, this argument has often been thought to conflict with central tenets of the Analytic and, more generally, the discursivity thesis. For the latter clearly entails that neither space nor time can be viewed as directly "given" to the mind in intuition as boundless or of infinite extent independently of the operation of the understanding. To be sure,  the account in the Aesthetic precedes the official introduction of this thesis at the beginning of the Analytic, and we cannot blame Kant for being unable to say everything at once. Nevertheless, any viable interpretation of the present claim must at least be compatible with this central Kantian thesis. And herein lies the problem. 

In addition to this general worry, there are at least two specific points at which this account appears to conflict with doctrines spelled out in the Analytic. The first concerns the Axioms of Intuition, where Kant claims that space can be represented only by means of a successive synthesis, which seems incompatible with its presumed givenness.<sup>39</sup> The second is Kant's enigmatic assertion at the end of the Analytic that "pure space and pure time . . . are to be sure something, as the forms for intuiting, but are not in themselves objects


that are intuited (*ens imaginarium*)” (A291 / B347). If space and time are not given as *objects* of intuition, in what sense can they be said to be given (or intuited) at all? 

In dealing with this issue, it is essential to return to the formula *intuitus, quem sequitur conceptus*. As I have already indicated, this expresses the thought that the conceptualization of space presupposes a preconceptual pattern or order (in Kant’s terms, a “pure manifold”), which both guides and constrains this activity. In virtue of its relation to the conceptual process, this manifold may be said to confront thought as a brute datum and, therefore, as something simply “given,” though not as a distinct object that might somehow be inspected independently of all conceptualization. Moreover, it follows from this that the latter is necessary for human cognition (and thus *a priori*), but not logically necessary.<sup>40</sup> With respect to the geometrical dimension of the issue, the basic idea was already articulated by Schultz, who wrote:

If I should draw a line from one point to another, I must already have a space in which I can draw it. And if I am to be able to continue drawing it as long as I wish, without end, then this space must already be given to me as an unlimited one, that is, as an infinite one. Correlatively, I cannot successively generate any cylinder or body except in space, that is to say, I can do so only because this space is already given, together with its quality which allows me to suppose that points are everywhere, and which enables me to generate, without end, the three dimensions of extension.<sup>41</sup>

Arthur Melnick has amplified this point to include the empirical representation of space. Concerned with the apparent contradiction between Kant’s infinity claim and certain doctrines of the Analytic, Melnick writes:

We do not perceive spatial regions (extents of objects in space) that are limitless or without bounds. Rather, we perceive space under the pre-conception (or, better, under the “*pre-intuition*”) that the bounded spatial extents we do perceive are parts of a limitless or unbounded space.<sup>42</sup>

The expression “pre-intuition” is appropriate here because it underscores Kant’s point that every determinate space is represented as a region or determination of the one unbounded space. This may be taken as a deep phenomenological point about the “form” or nature of our spatial (and temporal) experience.<sup>43</sup> This one unbounded space is “pre-intuited” in the sense that it is given together with every determinate space as the latter’s “horizon,” without itself being actually intuited as an object. Moreover, if this is correct, it follows that Kant’s controversial thesis that “space is represented as an infinite given magnitude” is best taken as a claim about the *a priori* structure of our spatial experi- 

ence in general, rather than as one about a unique (pre-conceptual) representation of an infinite space.

An interesting variation on this theme has been developed by Lorne Falkenstein, under the rubric of Kant's "formal intuitionism." Approaching the problem in information-processing terms, Falkenstein understands "intuitionism" to be the view that "a certain output is already contained in the input to a processor, so that it does not require any process (other than transmission or attention) to become known."<sup>44</sup> This is contrasted with "constructivism," which holds that "a given output is *not* already contained in the input to the cognitive system, but needs to be worked up out of that input by some process such as association, inference, comparison, combination, or composition."<sup>45</sup> By "formal intuitionism" Falkenstein understands a variation of the former in which the "input" includes not simply sensory content (sensations) but also the order or form in which this content is received.<sup>46</sup> For Falkenstein, then, Kant is a formal intuitionist because he maintains such a view regarding the given.

Expressed in more Kantian terms, this means that receptivity is taken seriously and given an essential cognitive function in structuring the form of our experience. Not only does it provide the data without which thought would have no content; it also presents the data in a certain fixed manner, which is independent of the conceptual activity of the understanding. Thus, on this view, the understanding (or imagination) does not produce a spatiotemporal order through its activity but merely uncovers or brings to consciousness one that is given independently of it, though not, as we shall see, independently of the nature of human sensibility.

In the light of this, the apparent contradiction between the Aesthetic and the Axioms is easily resolved. There is no conflict because Kant is concerned in the Axioms with the representation of determinate spaces and with the connection between such representation and the intuition of objects in them. The claim is that this representation presupposes a successive synthesis and that every determinate space must be conceived of as the product of such a synthesis, that is, as an extensive magnitude. Clearly, this is a claim about the conceptual conditions (rules) under which it is possible to represent a determinate space. Just as clearly, however, such a successive synthesis presupposes the givenness of the single, all-inclusive space that is to be determined.

Finally, this puts us in a position to consider (at least in a preliminary way) the notorious footnote in the B-Deduction, where Kant deals explicitly with the problem of the connection between the "original," "given," and "unlimited" space and time and determinate representations of spaces and times.<sup>47</sup> Al-

though this note is attached to the discussion of the synthesis of apprehension, which is the empirical synthesis that Kant contends is involved in sense perception, it is intended to explicate the claim made in the text that space and time are not only *a priori* forms of intuition but also themselves *a priori* intuitions with a manifold of their own. In an effort to explain this, Kant remarks:

Space, represented as **object** (as is really required in geometry), contains more than the mere form of intuition, namely the **comprehension** [*Zusammenfassung*] of the manifold given in accordance with the form of sensibility in an **intuitive** representation, so that the **form of intuition** merely gives the manifold, but the **formal intuition** gives unity of the representation. In the Aesthetic I ascribed this unity merely to sensibility, only in order to note that it precedes all concepts, though to be sure it presupposes a synthesis, which does not belong to the senses but through which all concepts of space and time first become possible. For since through it (as the understanding determines the sensibility) space and time are first given as intuitions, the unity of this *a priori* intuition belongs to space and time, and not to the concept of the understanding (§24). [B160–61, note]



We shall return to this note in chapter 7. For the present, our concern is with the contrast between a ‘form of intuition’ and a ‘formal intuition’, both of which fall under the label ‘pure intuition’. This contrast reflects, at the level of pure intuition, the general distinction made in the preceding chapter between an indeterminate (unconceptualized) and a determinate (conceptualized) intuition. Although Kant’s failure to introduce this distinction in the Transcendental Aesthetic has been the source of a good deal of the confusion concerning his analysis, this note makes it apparent that he thought it important to interpret the argument of the Aesthetic in the light of this distinction.

Once again, however, the situation is more complex than it initially appears. For if we apply the analysis of ‘intuition’ I sketched in the last chapter to the case of ‘pure intuition’, we are forced to distinguish three senses of the term. Not only must we contrast a ‘form of intuition’ (indeterminate pure intuition) with a ‘formal intuition’ (determinate pure intuition); we must also distinguish two senses of the former. This can mean either the form or manner [*Art*] of *intuiting*, which may be characterized as an innate capacity or disposition to intuit things in a certain way, for example, spatially and temporally, or the form (in the sense of essential order or pattern) of that which is *intuited*.<sup>48</sup>

The notion of a form of the intuited, distinct from both a form of intuiting and a formal intuition, is required to characterize the single, and all-inclusive, space that contains within it the manifold of spaces. Clearly, this space, which is the main focus of Kant’s analysis in the Transcendental Aesthetic, can be de-

scribed neither as a mere capacity to intuit nor as a formal intuition.<sup>49</sup> As the preceding analysis indicates, it must rather be construed as the “pre-intuited” framework that conditions and is presupposed by the actual representation of regions or configurations of space.<sup>50</sup> This is what Gerd Buchdahl refers to as “indeterminate space” or “spatiality.”<sup>51</sup> We shall see that it is space so construed that Kant claims to be transcendently ideal.

Finally, by ‘formal intuition’ is meant a determinate intuitive representation of certain “formal,” that is, universal and necessary, features of objects qua intuited. The crucial point here is that, as determinate, a formal intuition is a hybrid, requiring both the form of the intuited and a concept by means of which this form is determined in a certain way. A spatial formal intuition, with which the geometer is concerned, is the intuitive representation of the form or essential properties of the figure corresponding to a given geometrical concept. Such representations are the products of mathematical construction, which is itself ultimately governed by the given nature of space as the form of the intuited. In other words, this given nature, rather than merely the laws of logic, determines what is geometrically possible. This is also why Kant contends that geometry is both synthetic and *a priori*.

#### **D. The Transcendental Exposition (The “Argument from Geometry”)**

We must now consider the relevance of Kant’s view of geometry to the overall argument for the transcendental ideality of space. The so-called argument from geometry, which in the second edition is contained in the transcendental exposition, is frequently viewed as the main, if not the sole, support for the ideality thesis. But since Kant’s conception of geometry is generally thought to be untenable from a modern point of view, this makes it easy to dismiss out of hand the argument for transcendental idealism.

The concern of this section is to show that the centrality attributed to geometry by interpreters of Kant’s argument for transcendental idealism is misguided.<sup>52</sup> It is granted both that Kant advanced an argument from geometry to the transcendental ideality of space in the Transcendental Exposition and that his conception of space is intimately connected with his views on geometry.<sup>53</sup> What is denied is merely that Kant’s doctrine of the ideality of space is logically dependent on the latter.<sup>54</sup> Since nothing here turns on the merits or defects of Kant’s view of geometry, we may be quite brief.

As I have already indicated, the official task of a transcendental exposition is to show how an *a priori* representation can ground a certain body of synthetic

*a priori* cognitions. This requires showing, first, that the cognitions in question “actually flow from the given concept” and, second, “that these cognitions are possible only on the presupposition of a given way of explaining this concept” (B40). In other words, a transcendental exposition is designed to show that a given body of synthetic *a priori* knowledge (*P*) is possible only if there is a representation (*Q*) with certain specified properties. *Q* is thus a necessary condition for *P*, or equivalently,  $P \rightarrow Q$ .

This is precisely the connection that Kant asserts between geometry and the representation of space. It is assumed that “geometry is a science that determines the properties of space synthetically and yet *a priori*” and the question is “What, then, must our representation of space be for such a cognition of it to be possible?” (B40). It should be noted that the question concerns our *representation* of space, not space itself. Not surprisingly, Kant maintains that this representation must be both an intuition and *a priori*. Geometry is thus linked with the analysis of the representation of space, without a word being said about the ideality of space itself.

At this point, however, Kant suddenly asks: “Now how can an outer intuition inhabit the mind that precedes the objects themselves, and in which the concept of the latter can be determined *a priori*?” And in response he asserts,

Obviously not otherwise than insofar as it has its seat merely in the subject, as its formal constitution for being affected by objects and thereby acquiring immediate representation, i.e., intuition, of them, thus only as form of outer sense in general. [B41]

Here Kant does seem to indicate that the transcendental ideality of space is derived from the analysis of geometry. The argument, such as it is, consists of two steps. The first and previously noted step is the assertion that the *a priori* and intuitive character of the representation of space is a necessary condition of the possibility of geometry. The second is the claim that this entails that space itself must be a form of outer sense. A similar logical structure is also found in other texts, where Kant argues from the synthetic and *a priori* nature of geometry to the transcendental ideality of space. In each case, however, the move is mediated by an appeal to the *a priori* and intuitive character of the representation of space.<sup>55</sup>

Two important results follow from this. First, the transcendental ideality of space is only a necessary and not also a sufficient condition of geometry, construed as a synthetic *a priori* science of space. Consequently, the denial of the latter does not entail the denial of the former.<sup>56</sup> Second, the argument from geometry only moves to ideality by way of an appeal to the *a priori* and intuitive



character of the representation of space. As a result, if this can be established independently, then the ideality argument can proceed without any appeal to geometry. But the whole point of the Metaphysical Exposition is to show that the representation of space has just this character. It follows, therefore, that the argument for ideality can bypass completely the Transcendental Exposition or any considerations about the nature of geometry.<sup>57</sup> In fact, the most that any such considerations could provide is independent support for the contention that the representation of space is *a priori* and intuitive, which still leaves the task of demonstrating that space itself is transcendently ideal.<sup>58</sup> Moreover, the same thing may be said, *mutatis mutandis*, regarding the argument from the so-called incongruent counterparts, to which Kant sometimes appealed in support of the transcendental ideality of space.<sup>59</sup>



## II. THE IDEALITY ARGUMENT

Kant's real argument for the ideality of space is contained in the conclusions that follow immediately upon the Metaphysical and Transcendental Expositions. Kant draws two explicit conclusions regarding the content of the representation of space and then proceeds to claim that space itself is empirically real and transcendently ideal. Thus, it is here that the crucial move is made from a consideration of the representation of space to the nature of space itself. This turns out to be less of a transition that one might suppose, however, since Kant's basic claim is that space is nothing apart from our representation. The treatment of time has a parallel structure, though Kant adds that time, as the form of inner sense, is a formal, *a priori* condition of all appearances, since all appearances, as modifications of the mind, belong to inner sense (A34 / B50). Although this is of great significance for the argument of the Transcendental Analytic, because it is not directly relevant to our present concern we need not consider it here.

### A. Kant's Conclusions

The first conclusion is expressed in negative terms, stipulating what is *not* contained in our representation. Specifically, we are told that "[s]pace represents no property at all of things in themselves nor any relation of them to each other, i.e., no determination of them that attaches to objects themselves and that would remain even if one were to abstract from all subjective conditions of intuition" (A26 / B42). Since, apart from in the Preface, this is the first reference

in the *Critique* to things in themselves, the uninstructed reader is not in a position to interpret it. Nevertheless, Kant provides an important clue when he indicates that by a property of such things is to be understood a determination that attaches to them independently of the subjective conditions of intuition. This suggests that by 'things in themselves' is meant things considered apart (or in abstraction) from their relation to the subjective conditions of intuition. The claim is, therefore, that the representation of space does not contain any properties (including relational properties) that can be predicated of things qua considered in this manner.

The second, positively expressed conclusion asserts that "[s]pace is nothing other than merely [*nichts anders als*] the form of all appearances of outer sense, i.e., the subjective condition of sensibility, under which alone outer intuition is possible for us" (A26 / B42). Although there is a brief reference to the distinction between the matter and form of appearance at the beginning of the Aesthetic, the reader is once again not prepared for this conclusion. And, once again, Kant endeavors to clarify matters in an explanatory clause, where he indicates that by the "form of all appearances of outer sense" is meant a subjective condition of outer intuition.<sup>60</sup> Accordingly, Kant's claim is that the *content* of the representation of space, that is, what is actually represented (or, better, presented) therein, is only a subjective condition of human sensibility. The remainder of the paragraph then affirms the favorable consequences of this conclusion for the understanding of the possibility of *a priori* knowledge.

Finally, on the basis of these conclusions, Kant asserts:

We can accordingly speak of space, extended beings, and so on, only from the human standpoint. If we depart from the subjective condition under which alone we can acquire outer intuition, namely that through which we may be affected by objects, then the representation of space signifies nothing at all. This predicate is attributed to things only in so far as they appear to us, i.e., are objects of sensibility. [A26–27 / B42–43]

It is here and in the remainder of this paragraph that Kant affirms the transcendental ideality of space, as well as the compatibility of this ideality with its empirical reality. In essence, the ideality thesis is that spatial predicates are limited to "objects of sensibility," that is, appearances, or, equivalently, that these predicates are not applicable to things "when they are considered in themselves through reason, i.e., without taking account of the constitution of our sensibility" (A28 / B44). Correlatively, the empirical reality thesis is that these predi-

cates are applicable to outer appearances, which is equivalent to the assertion of the “reality (i.e., objective validity) of space in regard to everything that can come before us externally as object” (A28 / B44).

Apart from terminological matters, there are two major issues here: (1) What is Kant trying to prove? (2) How does he set about doing it? On the traditional reading, the answer to the first question is obvious: Kant is attempting to determine the ontological status of space (and time). On this reading, then, the first conclusion eliminates the standard ontological options and the second affirms the critical alternative as the only remaining possibility. In support of the latter, it is further claimed (on this reading) that the ontological status assigned to space as form of human sensibility does not threaten its empirical reality.

If the analysis of transcendental idealism presented here is correct, this reading is deeply misguided.<sup>61</sup> Clearly, the two conclusions are related as negative and positive, with the first dismissing the ontological options supposedly available to transcendental realism (basically the Newtonian and Leibnizian views), and the second affirming the doctrine of transcendental idealism. Once again, however, this need not be taken to mean that Kant’s positive conclusion should be understood as the successful ontological rival to the views he has repudiated. Rather than making an original move within the same game (traditional ontology), Kant might be taken as introducing a whole new game (transcendental philosophy).

In order to evaluate this possibility, it is necessary to consider briefly the nature of the game being abandoned. As we shall see in more detail in the fourth part of this book, Kant followed the metaphysical tradition (including his predecessors Wolff and Baumgarten) in understanding the object of ontology to be “things in general.” In other words, it is “general metaphysics” as contrasted with the various branches of “special metaphysics” (rational psychology, cosmology, and theology), each of which is concerned with a distinct transcendent object. Expressed in Aristotelian terms, it is the theory of “being qua being.”

Since, as we shall also see in more detail in later chapters, Kant explicitly denied that he was engaged in such a project in the *Analytic* (A247 / B303), it seems reasonable to assume that he would say this of the *Aesthetic* as well. Seen within the framework of the metaphysical tradition Kant inherited, this means that spatiotemporal properties and relations are not predicable of things or objects in general. This is the basic claim from which their inapplicability to things as they are in themselves follows as a corollary, since whatever holds of

things in general (considered merely *qua* things) must also hold of things considered as they are in themselves.

Consequently, to deny that Kant is providing an alternative ontology of space is simply to point out the obvious, namely, that he is denying that spatiality may be attributed to things in general (whether as property or relation). Such an attribution is explicitly rejected with the thesis that space is a form of human sensibility. Admittedly, one might claim in response that this denial of ontological standing is still a move within ontology. Doing so, however, is seriously misleading, since it suggests that it is a claim about how things “really are” *an sich*. In other words, it invites us to take Kant to be asserting that things are not really spatial but merely seem to us to be so. And this, as we have seen, engenders all of the difficulties concerning transcendental idealism, traditionally understood as a metaphysical thesis.

The issue may be further illuminated by a return to the notion of “warranted assertibility” introduced in chapter 2. So construed, Kant’s empirical realism may be viewed normatively as a “warrant” for attributing spatiality (magnitude, shape, and location) *a priori* to all objects of outer sense. It is, however, a “restricted warrant,” with the restriction being to objects of human experience, as opposed to an “unrestricted warrant” to apply them to things in general. Moreover, the grounds for this restriction are normative, since they stem from the conditions under which the warrant is issued in the first place, namely, that space has been shown to be a condition of outer intuition.

That Kant saw himself as doing something like this is suggested by his striking claim that it is only from the “human standpoint” that we can speak of space and that if we depart from this or, equivalently, the subjective condition of outer intuition, then the representation of space “signifies nothing at all.” To speak of space from the “human standpoint” is to consider it in terms of what in chapter 2 was described as the anthropocentric model of cognition, which is contrasted with the theocentric model of transcendental realism. Accordingly, to claim that the representation signifies nothing at all if we depart from that standpoint or, what here amounts to the same thing, the “subjective conditions of outer intuition,” is to claim that it is a mistake to think that space has an *an sich* reality of *any* sort. In a word, it is to “deontologize” space. Once again, however, this does not mean that things only *seem* to us to be spatial because they are perceived through the distorting medium of outer sense. It is rather that things *really are* spatial in the only meaningful sense in which this may be claimed, namely, considered as objects of possible experience.

Nevertheless, even if this is accepted as a statement of what Kant is trying to demonstrate, we are left with the question of how (and whether) he demonstrates it. In short, we need an argument showing us why we should give up our common way of thinking about space as an entity, property, or objective relation of some sort and reconceptualize it in the radically new manner Kant suggests. No such argument seems forthcoming, however, which is precisely what has led so many interpreters to conclude that Kant's real argument must be based on the synthetic *a priori* character of geometry. Moreover, many interpreters who do find an argument here that is independent of an appeal to geometry see it as resting on nothing more than the general connection between apriority and subjectivity.<sup>62</sup> But such a line of argument has two defects. First, it renders the whole discussion of the intuitive nature of the representation of space idle in the ideality argument. Second, it is incapable of explaining how the subjective origin of the representation of space can justify the claim of the transcendental ideality of space itself.

## B. In Search of an Argument

The *Prolegomena* contains important clues for the reconstruction of the requisite argument. After noting that the possibility of mathematics rests upon *a priori* intuition, Kant raises the question of the possibility of intuiting something *a priori*. Significantly, this problem does not arise for concepts, at least not in the same way. Kant writes: "Concepts are indeed of the kind that we can quite well form some for ourselves *a priori* (namely, those that contain only the thinking of an object in general) without our being in an immediate relation to the object" (Pro 4: 282; 78). Kant's point is that since concepts never relate immediately to objects, they can be formed independently of any experience of them. Thus, even though they would be empty, it is possible to form concepts for which no corresponding object can be given. But since an intuition relates immediately to its object, there could be no analogous empty intuition. It is the apparent incompatibility of this immediacy or presentational requirement with its presumed apriority that renders the notion of an *a priori* intuition problematic. This seems to require that an object somehow be given to the mind before it is actually experienced, which is a contradiction in terms. Accordingly, there is a need to explain "how . . . an *intuition* of the object [can] precede the object itself" (Pro 4: 482; 78).

The problem here concerns the content or intentional object of a putatively *a priori* intuition (its "objective reality" in the scholastic-Cartesian sense). In other words, it concerns the question: What is such an intuition an intuition

of? Kant notes that it cannot be an intuition of things as they are in themselves, which is merely a reiteration of the first conclusion of the *Critique*. He now goes further, however, and suggests that even an empirical intuition would be impossible on this assumption. As he puts it, “[I]t is incomprehensible how the intuition of a thing that is present should allow me to cognize it the way it is in itself, since its properties cannot migrate over into [*hinüber wandern*] my power of representation” (Pro 4: 282; 78). Although Kant’s manner of expression is somewhat fanciful, he appears to be alluding to the scholastic theory of perception (the doctrine of intentional species), which was replaced by the modern “way of ideas” for roughly the reasons Kant suggests.

Kant’s real concern is with *a priori*, not empirical, intuition, however, and the problem is to determine what such an intuition could contain or present to the mind. Having rejected the view that it represented things as they are in themselves as incompatible with this assumption, he concludes:

There is therefore only one way possible for my intuition to precede the actuality of the object and occur as an *a priori* cognition, *namely if it contains nothing else except the form of sensibility, which in me as subject precedes all actual impressions through which I am affected by objects*. [Pro 4: 282; 78]

The claim is thus that an *a priori* intuition is possible if and only if it contains or presents to the mind a form of its own sensibility. Admittedly, this seems mysterious, since there is no explanation of what it might mean to intuit a form of sensibility. Nevertheless, by avoiding any specific reference to either space or time, the putative forms of human sensibility, this formulation has the virtue of generality, which helps to clarify the conceptual nature of the issues involved. The argument consists of two steps, corresponding to the two conclusions in the *Critique*.<sup>63</sup> The first maintains that an *a priori* intuition is *not possible* if its content is a determination (whether intrinsic or relational) of things as they are in themselves. The second affirms that such an intuition *is possible* if its content is a form of sensibility. Assuming the exhaustiveness of these alternatives and the results of the Expositions, it follows that space and time, the contents of these representations, are nothing but forms of sensibility.

*The “Not If” Portion of the Argument.* As already indicated, the first conclusion rules out the possibility that space could be either a property or relation of things as they are in themselves because neither of these could be intuited *a priori*. Although these correspond to the Newtonian and Leibnizian positions, respectively, we have seen that Kant’s target is *all* transcendently realistic theo-

ries of space (or time), of which Newton's and Leibniz's are simply the most important. Thus, even though Kant is concerned in the Aesthetic (and elsewhere) with the relative advantages and disadvantages of these views (which he terms those of the "mathematical" and the "metaphysical students of nature"), his core ideality argument turns rather on what they have in common, namely, the assumption that space and time are determinations (of whatever kind) of things as they are in themselves.

Approaching the argument in this way makes it possible to deal succinctly with the otherwise vexing question of Kant's elimination of the Newtonian alternative. This question is vexing because, as commentators have noted, the Metaphysical Expositions seem directed mainly against the Leibnizian-relational view and leave the Newtonian-absolute view virtually untouched.<sup>64</sup> Accordingly, it might seem that Kant must appeal to additional considerations, including perhaps its unappealing metaphysical and theological implications, in order to eliminate this alternative and complete the argument.<sup>65</sup> But in that case, the conclusions would not follow from the Expositions alone.

If we follow the suggestion of the *Prolegomena*, however, this common way of analyzing the argument of the *Critique* can be seen to be mistaken.<sup>66</sup> For the main lesson to be learned from that is that whether space be taken to represent a property or a relation of things in themselves is a matter of relative indifference, since neither could be intuited *a priori*. Accordingly, the first or "not if" portion of Kant's argument turns entirely on a claim about the conditions of the possibility of intuiting something *a priori*.<sup>67</sup>

So construed, the argument is no stronger than its premise that we have an *a priori* intuition of space. Moreover, this is a premise that the Leibnizian at least would deny. Indeed, given the problematic nature of an *a priori* intuition, it might be viewed as a virtue that a position is unencumbered with any such conception.

Since it emphasizes synthetic *a priori* knowledge, the most obvious response is to appeal to the Transcendental Exposition. But since the latter itself turns on the assumption that the original representation of space is both *a priori* and intuitive, we are inevitably led back to the Metaphysical Exposition as the ultimate basis for Kant's conclusion. Thus, again confining ourselves to the treatment of space, a brief review of the results of this Exposition seems to be in order.

As we have seen, these results are twofold. First, the representation of space was shown to have a complex epistemic function with respect to outer appearances (whence its apriority). More specifically, it was shown to ground the pos-

sibility of outer intuition, both in the sense of providing a condition of the perception of the relations of outer appearances to the self and to each other (the first argument) and in the sense of being a condition of the very givenness of such appearances (the second argument). Second, it was shown to fulfill this function in a certain manner, namely, by providing an infinite or boundless and homogeneous framework in which these appearances are given and related to one another, rather than a generic concept under which they fall (whence its intuitive nature).

Consequently, the question becomes whether the representation of space could function in this way unless it were both intuitive and *a priori*, and the answer is clearly negative. First, given Kant's understanding of the concept-intuition distinction, that within and respect to which outer appearances are related cannot be regarded as a concept under which they fall and through which they are recognized as such. Second, and more controversially, there is a contradiction involved in the assumption that the representation of something that supposedly functions as a condition of the possibility of the experience of objects could have its source in the experience of these very objects. This is contradictory because it entails that experience is possible apart from something that is stipulated to be a condition of its possibility. Thus, if the representation of space plays the role it is claimed to in the *Metaphysical Exposition*, it must be *a priori* and, as such, have its ground in the cognizing subject rather than in the nature of the objects as they are in themselves.

Nevertheless, it is important to keep in mind the limited nature of this result. Of itself, it neither establishes anything positive about the nature of space nor even rules out the possibility that space and time might, as Newton and Leibniz maintain, be either properties or relations of things in themselves. Rather, it shows merely that no such transcendently realistic space could provide the content of our representation and, therefore, be the space referred to in either empirical judgments or mathematics. Consequently, if Kant is to move beyond this negative result, he must show that his conception of space as a form of human sensibility is capable of explaining how space could be intuited *a priori*.

*The "If" Portion of the Argument.* This is the task of the second part of the argument, which asserts that an *a priori* intuition is possible if it contains or presents to the mind its own form of sensibility. Like the first conclusion, this is perfectly general and does not involve any specific references to the representations of space or time as pure intuitions. Accordingly, the issues it poses are largely



terminological: for example, the meaning of 'form of sensibility' and whether such a form is the sort of thing that could be intuited. Unfortunately, Kant's own definitions at the beginning of the *Transcendental Aesthetic*, where one would expect to find answers to these questions, are not very helpful. As H. J. Paton points out, Kant tends to treat 'form of appearance', 'form of intuition', 'form of sensibility', and even 'pure intuition' as if they were virtually equivalent expressions, which seems to make the contention that a pure intuition is, or has as its content, a form of sensibility almost a matter of definition.<sup>68</sup>

In order to see that this is not really the case, it is necessary to consider more closely some of these definitions. Since it is in many ways the most basic, we shall begin with the expression 'form of appearances'. If we are to avoid begging the question, 'appearances' must be taken here in a neutral sense, that is, as not already implying any ideality. Instead, it refers to the objects that are actually given in experience, in contrast to those that are merely conceived. Correlatively, 'form' must be taken to mean condition and 'matter' that which is conditioned or determined by the form.<sup>69</sup> In the second edition, Kant defines the form of appearance as "that which allows the manifold of appearance to be ordered in certain relations" (B34). Otherwise expressed, a form of appearance is a feature of the appearance in virtue of which its elements are orderable or relatable to one another in intuition. Thus, as we saw in chapter 2, it is not itself an order (as space and time are for Leibniz) but a condition or framework of such an order. The first apriority argument, in particular, maintains that the representation of space functions in this way.

We have already considered the expression 'form of intuition' and its inherent ambiguity. We saw that it can designate either the formal features or structure of the objects intuited or the mode or manner (*Art*) of intuiting them. In the former case, it is equivalent to 'form of appearances' and, therefore, ontologically neutral. In the latter case, however, it involves an explicit reference to mind. For a form of intuiting is a characteristic of mind and, more particularly, of its receptive capacity, not of things as they may be apart from their relation to mind. Like 'form of intuition', 'form of sensibility' can be taken in two senses. But unlike the former expression, both of these senses involve a reference to mind. More precisely, 'form of sensibility' can designate either a form of sensibly intuiting, which Kant sometimes also terms a 'form of receptivity', or a form of objects qua sensibly intuited. Henceforth, these will be referred to as "form of sensibility<sub>1</sub>" and "form of sensibility<sub>2</sub>," respectively. The main point is that, in claiming that a form of appearances is a form of sensibility<sub>2</sub>, one is also claiming that it is a form that pertains to these appearances (given objects) only

in virtue of the subjective constitution of the mind, that is, its form of sensibility<sub>1</sub>, which is just what Kant's conclusions are supposed to establish.

Given these preliminaries, we are now in a position to formulate the "if" portion of the argument concisely. It maintains that if the content of a given intuition is a form or formal feature of objects of intuition (the *intuited*) that pertains to these objects only in virtue of the constitution of the mind (its form of *intuiting*), then that intuition must be *a priori*. This is because, first, the content of such an intuition would be universal and necessary (at least for all subjects equipped with the same form of intuiting) and, second, its source would lie neither in the objects themselves nor in any sensible data (sensations) produced by the affection of the mind by these objects. For the second reason it would also be "pure," that is, independent of sensation. Once again, this claim is completely general, applying to pure intuition as such, without any specific reference to Kant's analysis of the representation of space. Nevertheless, it does show that we can account for the possibility of an *a priori* intuition of space, if we assume that it is (or contains) a form of sensibility<sub>2</sub>.

Do these conclusions jointly entail that space is a form of sensibility<sub>2</sub> and, as such, transcendently ideal? Assuming that Kant's definitions are in order and that the claims of the Metaphysical Exposition are cogent, the issue (as it so often does) turns on the exhaustiveness of the alternatives presented and supposedly eliminated. The problem seems exacerbated, however, by an apparent gap in Kant's argument. The first step purports to show merely that an *a priori* intuition is not possible if its content or intentional object is either a property or relation of things as they are in themselves. But this seems to leave in place at least two possibilities: that the content of this representation might be something other than the three (or four) alternatives suggested; and that, even though the content of our *representation* of space is a form of sensibility, *space itself* might nonetheless be either a property or relation of things as they are in themselves and, therefore, transcendently real. The latter is the topic of the next section, but it is necessary to say a word about the former at this point.

As I have already indicated, the first part of the argument turns on the shared transcendently realistic assumption of the rejected alternatives. In other words, in spite of the deep differences between them, from Kant's transcendental point of view both Newton and Leibniz assume that the question of the nature of space and time must be addressed at the ontological level as a matter of a property or relation of things as they "really are," seen from a God's-eye view. In short, theirs is a family quarrel. But since it is precisely this assumption that

makes it impossible to account for the possibility of their being intuited *a priori* (as is required by the Metaphysical and Transcendental Expositions), the elimination of these alternatives extends to all accounts sharing this assumption, that is, all transcendently realistic accounts. Moreover, since we have seen that transcendental realism and transcendental idealism are mutually exclusive and exhaustive metaphilosophical alternatives, it follows that the elimination of the former leaves only the latter. Thus, if the latter alternative can account for the possibility in question (which the second part of Kant's argument shows to be the case), then it must be accepted as the only viable alternative.

The argument so construed is neither question begging nor, as is sometimes claimed, based on a terminological slight of hand, which effectively trivializes it by making ideality largely a matter of definition. It is not question begging because, even though it presupposes the dichotomy between transcendental realism and transcendental idealism in its framing of the problem, it does not pre-judge the issue in favor of the latter. It does not trivialize the issue because it does not make the ideality of space follow merely from the definitions. Although we have seen that these definitions are crucial for the understanding of Kant's thesis, we have also seen that they serve largely to clarify the results of the Metaphysical Exposition, where most of the "heavy lifting" is accomplished.

### C. The Problem of the Neglected Alternative

One of the classical objections to Kant's ideality argument is the so-called neglected alternative. This objection can be traced back to Kant's own contemporaries and has been formulated in a variety of ways.<sup>70</sup> The basic charge is that even if (for the sake of argument) we grant that space *as we represent it* is a form of human sensibility, this does not remove the possibility that space itself does not *also* have an *an sich* status (whether as property or relation). Expressed in terms previously introduced, the fact that the spatiality warrant does not extend to things as they are in themselves does not mean that things, so construed, are not spatial. It shows only that claims about their spatiality are not "warrantably assertible." Indeed, it is often suggested that Kant cannot deny such a possibility without contradicting the critical principle that things as they are in themselves are unknowable. The objection is succinctly expressed by Kemp Smith, who writes:

Kant recognizes only two alternatives, either space as objective is known *a posteriori*, or being an *a priori* representation it is subjective in origin. There exists a third alternative, namely that although our representation of space is subjective in origin, space is itself an inherent property of things in themselves.<sup>71</sup>

In order to assess this objection, it is necessary to specify the nature of the alternative that is supposedly neglected. Three possibilities suggest themselves, which listed in order of decreasing strength are: there is a numerical identity between the space that is a form of sensibility and a space pertaining to things as they are in themselves; these two spaces are qualitatively, though not numerically, identical, that is, they share a common structure; and there is a similarity or analogy, though not an identity, between the two spaces. Since these are significantly different alternatives, each requires separate discussion.

The first and strongest of these is easy to dismiss. As Reinhold had already pointed out in his analysis of the concept of representation and the form thereof, the form of a representation is precisely what makes it a representation and, therefore, distinguishes it from anything else.<sup>72</sup> Consequently, assuming a parallelism between "representation talk" and "sensibility talk" on this point, it would be absurd to suggest that a form of sensibility could be numerically identical with (the very same thing as) a property or relation of things as they are in themselves.

This does not get us very far, however, since it only rules out the neglected alternative in its least plausible form. As Vaihinger noted, the neglected alternative that Kant's critics had in mind is not that a form of representation (sensibility) is *also* a form of things as they are in themselves but rather that there is something in the things as they are in themselves corresponding to this form of representation.<sup>73</sup> Whether this correspondence is understood as a qualitative identity or merely a similarity or analogy remains undetermined.

Assuming once again a parallel between Kant's 'form of sensibility' and Reinhold's 'form of representation', Reinhold's analysis can be extended to address at least the first and stronger of these alternatives. A 'form of representation' can designate either a form (mode or manner) of representing or a form of what is represented. Moreover, like Kant's form of sensibility,<sup>2</sup> the latter pertains to what is represented only in virtue of a specific mode or manner of representing. In short, a reference to mind and its capacities is built into the very notion of such a form, just as a reference to a sensibly affected or receptive mind is built into the Kantian conception.

It is conceded by the objector, however, that space is (at least) such a form. Consequently, it must likewise be granted that spatial predicates cannot be applied to things as they are in themselves. Or, more precisely, it must be conceded that such predicates would not be qualitatively identical to the spatial predicates connected with sensibility. Since mind independence is a defining characteristic of the former and mind dependence of the latter, speaking of a

qualitative identity here is like speaking of an identity between a sensation and something that is exactly like a sensation, except for the fact that it cannot be sensed.

This leaves us, then, with the third and weakest version of the allegedly neglected alternative: that there might be a similarity or analogy between our “human space” as a form of sensibility and a “real” space pertaining to things as they are in themselves. Because of its vagueness this possibility is much more difficult to dismiss. Indeed, it is far from clear that Kant could, or would even need to, do so.<sup>74</sup> Nevertheless, we cannot be satisfied by pointing to this vagueness, since there is a non-trivial form of the similarity thesis that is not ruled out by the preceding analysis and that seems to constitute a direct threat to the non-spatiality claim.

The problem is that the above argument only succeeded in ruling out the possibility of a qualitative identity by pointing to the necessary mind dependence of a form of sensibility. Accordingly, it leaves open the possibility that a transcendently real space might be exactly like the space of human sensibility *except for this feature*. But since this is arguably just what proponents of the neglected alternative had (or have) in mind, it also leaves that venerable objection virtually intact.<sup>75</sup>

As Falkenstein has pointed out, however, the Aesthetic provides materials for the rejection of the latter possibility, since it indicates that the representation of space contains a certain kind of order. Underlying this analysis is a distinction between two types of order: a “presentational order,” which, as the term suggests, is one in which the mind receives its data in intuition, and a “comparative order,” which involves properties or qualities, such as colors. Given this distinction, Falkenstein argues that spatial and temporal orders are of the former sort, whereas an order of things as they are in themselves would have to be a comparative order of internal properties (as, for example, in the Leibnizian monadology). Accordingly, the question becomes whether it is meaningful to affirm an isomorphism, or even a significant similarity, between two such radically distinct types of order. Falkenstein denies this on the grounds that the Expositions and Kant’s account of incongruent counterparts show that a presentational (spatiotemporal) order of phenomena is indifferent to the internal properties of the objects ordered therein, whereas a comparative order is not.<sup>76</sup>

Falkenstein’s analysis is basically correct. As forms of sensibility, space and time are forms or conditions of the order in which phenomena present themselves in intuition. Moreover, as such, they are quite distinct from any purely conceptual ordering that might pertain to things as they are in themselves (as

conceived by some pure understanding). Falkenstein denies, however, that this conclusion entails any idealistic commitments; and in his later treatment of the topic he denies that it suffices to establish “the strong thesis that space and time are ‘nothing’ outside of our experience.”<sup>77</sup> Accordingly, before concluding this discussion, we must briefly address these points.

Falkenstein denies that his conception of a presentational order entails idealism because of its empirical nature.<sup>78</sup> And he is correct in noting that, as given, a presentational order might reflect what is given to the mind in experience rather than its manner of receiving it (a form of sensibility<sub>1</sub>). In short, there is no direct inference from the concept of a presentational order to its mind dependence. This is beside the point, however, since his empiricist alternative is already ruled out by the argument of the Metaphysical Exposition. In other words, the concept of a presentational order does not serve to establish the subjective source of our representation of space; rather (assuming the latter), it explains why such an order cannot be replicated by an order of things as they are in themselves. These are quite distinct issues.

With regard to the second point, Falkenstein’s denial that the conception of a presentational order suffices to rule out the neglected alternative appears to be motivated by a worry about reducing the issue to a matter of definition (whereby ‘space’ is stipulatively identified with ‘presentation space’). Thus, while admitting that a spatial order of things in themselves must differ considerably from the “presentational space” of human experience, Falkenstein denies that this precludes the former order from being in *any sense* spatial. For all we could know, he suggests, it might be a “quality space, where the locations of things in themselves are determined by internal properties that shade off into one another in various dimensions.”<sup>79</sup>

There are two things to be said in response to this. First, Falkenstein’s worry about reducing the issue to a matter of definition, which is a version of the triviality objection, is misguided. Although it is a matter of stipulation to characterize the space of human experience as a “presentational space” or, in Kantian terms, a “form of sensibility,” that this space has the form of a presentational space is not merely stipulated but shown in the Metaphysical Exposition (particularly the intuition arguments). As for Falkenstein’s alternative “quality space,” if conceivable at all, it bears so tenuous a resemblance to the presentational space with which Kant was concerned as not to constitute a serious threat to the non-spatiality thesis. In short, it is an arbitrary, ad hoc hypothesis rather than a serious alternative that Kant may be said to have neglected.

Second, and most important, as is the case with all proponents of the ne-

glected alternative objection, Falkenstein's scruples about the ideality thesis rest upon a transcendently realistic analysis of the situation. That is to say, he takes it as an ontological thesis about the true nature of *an sich* reality. Moreover, taking it in this way, he is quite correct to raise apparently ignored metaphysical possibilities, such as that of a quality space. Once again, however, if the interpretation of transcendental idealism advocated here is correct, this misses the whole point. This idealism is not an ontological thesis about how things "really are" (non-spatial and non-temporal), when seen from a God's-eye view. It is rather a critical thesis about the conditions of the cognition of things viewed from the "human standpoint," which is the only standpoint available to us. One can, of course, quarrel with Kant's claims that space and time are such conditions. What one cannot do is claim that it is possible *both* for space and time to be such forms *and* for things as they are in themselves to be spatial or temporal in any meaningful sense. Not only do we have no warrant for this, we could not conceivably get one.

## Chapter 6 The Intellectual Conditions of Human Cognition: Kant's Metaphysical Deduction

By the “intellectual conditions of human cognition” is meant the pure concepts of the understanding. Following Aristotle, Kant also terms them categories. The demonstration of their status as necessary conditions of the possibility of experience is the task of the Transcendental Deduction and will be the concern of the next chapter. But before this can be undertaken, Kant has the preliminary, yet indispensable, task of showing that there are such concepts and of specifying them. This is the concern of “On the Clue to the Discovery of All Pure Concepts of the Understanding” and the subject of the present chapter. In the second edition, Kant also refers to this section as the “Metaphysical Deduction,” and he claims that in it “the *a priori* origin of the categories in general was established through their complete coincidence with the universal logical functions of thinking” (B159).<sup>1</sup>

In linking the categories with the logical functions of thinking, Kant attempted both to establish their credentials as the fundamental forms of discursive thought and to demonstrate the completeness of his list. The key to both parts of the project lies in the identification of discursive thinking with judging, which makes it possible to under-



stand the special status of these concepts by seeing their connection with the act of judgment, and to guarantee their completeness by providing an exhaustive account of the forms of judgment.

Unfortunately, in spite of the significance Kant attributes to it, the Metaphysical Deduction is among the most widely rejected parts of the *Critique*. Although the criticisms are varied, they focus largely on two main points. One concerns the completeness of the table of logical forms from which the categories are supposedly derived. Ironically, much the same criticism that Kant raised against Aristotle regarding the categories, namely, that he arrived at them in an ad hoc manner (A81 / B107), is applied to Kant's presentation of the forms of judgment on which his derivation of the categories is based. Thus, it is frequently charged that there is no systematic principle in terms of which the completeness of these forms can be understood; that Kant simply took them as he found them in the logic texts of his time; and, even worse, "doctored up" the table of logical forms to make it agree with his table of categories.<sup>2</sup>

The other major line of criticism concerns Kant's move from the forms of judgment to the categories. Here the objection is that Kant's attempt to derive categories or ontological concepts from logical forms is deeply misguided. Among contemporary critics, this charge is often justified by appealing to the development of modern logic. Regardless of how plausible Kant's project might appear within the context of classical Aristotelian logic, it is regarded as an obvious non-starter when viewed in the light of modern truth-functional and predicate logic.

Once again, Strawson may serve as spokesperson for this point of view. Appealing to the modern conception of "logical form," Strawson argues that, in order to yield a category, a logical form must express not simply a *possible* form but one that is necessary and fundamental. This reflects the view that the forms taken as primitive in a logical system are a matter of choice and that the only two ideas that are truly fundamental from the standpoint of modern logic are truth-functional composition and quantification. But, as Strawson points out, these "logical forms" are not likely sources of categories, since apart from the distinction between particulars and kinds, there are no ways in which we must conceive of objects in order to reason about them under these forms. Moreover, since the latter can already be inferred from the concept-intuition distinction, Strawson concludes that Kant's whole venture into the logical domain is beside the point.<sup>3</sup>

With these problems in mind, I have divided this chapter into three parts. The first discusses the completeness problem in the light of important recent

work on the topic.<sup>4</sup> The second and third address the derivation problem in two steps. First, basing the analysis on Kant's account of judgment in §9, it is argued that for Kant every form of judgment necessarily involves a certain mode or manner of conceptualization (logical function). Second, by examining the dense argument of §10, it is claimed that, while distinct from the categories, these functions provide the required "clue" to the discovery of the latter, because the two are isomorphic expressions of a single form of understanding operative in two domains (discursively in judgment and pre-discursively with reference to intuition).

### I. THE COMPLETENESS PROBLEM

In §9 Kant presents his cryptic account of what he terms the "logical function of the understanding in judgment." The underlying claim is, "If we abstract from all content of a judgment in general, and attend only to the mere form of the understanding in it, we find that the function of thinking in that can be brought under four titles, each of which contains under itself three moments" (A70 / B95). This is the notorious completeness thesis, which, as I have already noted, has been the subject of much dismissive criticism.

Before addressing the problem itself, we must consider the claim that the whole endeavor is misconceived, since Kant himself denied both the need for and the possibility of anything like a "completeness proof."<sup>5</sup> Indeed, support for such a claim seems to be provided by a number of texts, most notably the following well-known passage in the B-Deduction:

But for the peculiarity of our understanding, that it is able to bring about the unity of apperception *a priori* only by means of the categories and only through precisely this kind and number of them, a further ground may be offered just as little as one can be offered for why we have precisely these and no other functions for judgment or why space and time are the sole forms of our possible intuition. [B145–46]<sup>6</sup>

Nevertheless, there are at least two compelling reasons why this should not be read as ruling out the project of establishing the completeness of the table of logical functions. First, the text itself does not require such a conclusion. It speaks not of *any* ground but of any *further* ground.<sup>7</sup> In the case of the categories, with which Kant is primarily concerned in the passage, what is denied is the possibility of providing a further ground, beyond the one just offered, namely, the synthetic unity of apperception. In the case of the logical functions, the parallel claim would be that no ground can be provided beyond the analy-

sis of the functions essential to discursive cognition sketched in §9.<sup>8</sup> Second, Kant himself unequivocally indicates in several places that establishing the completeness of the table is essential to his project, since without it the determination of the categories and their systematic completeness (which is the main order of business) would be impossible.<sup>9</sup> Thus, whatever its inherent difficulties, the completeness problem must be addressed by any serious interpretation of the *Critique*.

Fortunately, the problem becomes considerably more tractable once we narrow its scope. Even though Kant refers simply to judgment and its forms, it is clear from his analysis that he has only certain judgment types in mind: those that determine an object through concepts, which Kant generally terms “judgments of the understanding,” or “logical judgments.”<sup>10</sup> In light of the crucial connection between judgment and discursive thought, these will be referred to here as “discursive judgments.” They include all judgments capable of serving as premises of syllogisms but exclude, for example, mathematical judgments, which are non-discursive for Kant since they rest on the construction of concepts rather than on concepts, as well as judgments involving indexicals and proper names.<sup>11</sup>

To clarify matters further, it is useful to break the completeness problem down into two parts corresponding to the two organizing principles of the table. One deals with the four titles under which the twelve functions are grouped; the other (and more difficult part) with the three moments falling under each title.

#### A. The Completeness of the Titles

The four titles or basic headings into which Kant divides judgments in terms of their form are: quantity, quality, relation, and modality. Apart from relation, which replaces the then standard division of judgments into simple and complex, there is nothing remarkable about this list; nor is there meant to be. As Kant himself notes, he found before him “already finished though not yet wholly free of defects, the work of the logicians” (Pro 4: 323; 115–16). What is new, however, is the systematic significance attributed to this set of headings and the claim that each of the logical functions necessarily falls under one of them.

Since Kant makes this claim without further argumentation and since it is hardly self-evident, we must assume that he took it to be a consequence of the account of judgment offered in the first section of the “Clue.” As we saw in chapter 4, the notion of function is central to this account. Indeed, within a single paragraph of perhaps unparalleled density Kant claims that “concepts

rest on functions” (in contrast to intuitions, which rest on affections); defines ‘function’ as the “unity of the action of ordering different representations under a common one”; characterizes all judgments as “functions of unity among our representations”; and suggests that the “functions of the understanding can . . . all be found together, if one can exhaustively exhibit the functions of unity in judgments” (A68–69 / B93–94). Moreover, the following section, which supposedly explains why an exhaustive exhibition is possible, is entitled “On the logical function of the understanding in judgments”; and its central claim is, “If we abstract from all content of a judgment in general and attend only to the mere form of the understanding in it, we find that the function of thinking can be brought under four titles, each of which contains under itself three moments” (A70 / B95).

Although each of these mentions of ‘function’ is significant, our present concern is largely with the last four. First, by the “logical function of the understanding in judgments” is to be understood its function as analyzed in general as opposed to transcendental logic. The logical function of the understanding is, therefore, to be contrasted with its putative “real function,” which at this stage of the analysis is very much in doubt and is, in any event, a concern of transcendental logic. Second, the “functions of the understanding,” which are what Kant is trying to determine, are just the logical functions of judgment, that is, forms of conceptualization operative in judgments of the corresponding form. Correlatively, the “functions of unity in judgment” are what is supposedly being exhibited in the table. Thus, if the project is to make sense these obviously must be distinguished from each other.<sup>12</sup> Third, what must be brought under the four titles are the functions of thinking in judgments, which suggests that these should be seen as specifications of the generic function of thinking as manifested in all discursive judgments, regardless of their content.

In contrast to the definition, all of these uses of ‘function’ appear to take the term in the familiar Aristotelian-physiological sense of task. Accordingly, the basic task or function of discursive thought is to judge, and Kant’s claim is that this may be broken down into a number of sub-functions that fall into certain types. Thus, the headings refer to the types of sub-function required for a successful exercise of the generic function. Appealing to a biological analogy, one might say that just as the function of the eye, namely, to see, may be broken down into several sub-functions, such as color, shape, and distance vision, so the function of the understanding, namely, to judge, may be broken down into four (and only four) types of sub-function: quantity, quality, relation, and modality.

In order to test the aptness of this analogy, let us briefly revisit Kant's functional account of judgment in the first section of the "Clue." As we saw in chapter 4, Kant's basic claim is that the essential task of a (discursive) judgment is to produce an object-related unity of representations under a concept. The underlying assumption, which is first made explicit in the B-Deduction (§19), is that only such a relation of representations is capable of a truth value. And given this result, it is relatively easy to see that every judgment necessarily involves at least three sub-functions corresponding to the first three of Kant's titles.<sup>13</sup>

Keeping to the order in Kant's table, the first of these sub-functions may be characterized as defining an extension. Thus, in Kant's model judgment "All bodies are divisible," the predicate divisibility is related to everything falling under the concept of 'body'. This is the function expressed in the title "quantity," and it is apparent that it is essential to the exercise of the generic function of judging, since without it, it would be impossible to determine the scope of the predicate of the judgment and, therefore, assess its truth value.

The second essential sub-function is that of making a claim (assertion or denial) on the basis of the determination of the scope of the predicate. This is the function expressed in the title "quality." It is likewise essential to judgment, since without it there would be nothing to assess with respect to truth value. Moreover, quality, so construed, clearly presupposes quantity, since the claim to be assessed necessarily involves a quantified subject (for example, all, some, or a single body).

This brings us to the function of relation. In the case of the simple subject-predicate judgment, the judgment clearly requires a determination of the relation between the two concepts being connected. And, as we shall see shortly, in complex judgments, whose elements are themselves judgments rather than concepts, what requires determination is the relation between the constituent judgments. In other words, since judgment involves a mediate representation of an object or a "representation of a representation" of it, it is necessary to determine which representation is mediated by which or, equivalently, which is the "condition" and which the "conditioned." This, then, is the judgmental function of relation, and it too is clearly essential to judgment. Moreover, just as quality presupposes quantity, the function of relation presupposes both, since these jointly provide the conditions of the determination of the relation expressed in the judgment.

If anything, however, this analysis appears to be *too successful*. The problem is that it seems to provide a complete account of the cognitive tasks requisite for a judgment like "All bodies are divisible," without any reference to modality. Ac-

cordingly, if the completeness claim is to be substantiated, the latter must somehow be brought into the story. Here the crucial text is Kant's remark,

The modality of judgments is a quite special function of them, which is distinctive in that it contributes nothing to the content of the judgment (for besides quantity, quality, and relation there is nothing more that constitutes the content of a judgment), but rather concerns only the value of the copula in relation to thinking in general. [A74 / B99–100]

At first glance it may seem peculiar for Kant to describe quantity, quality, and relation, which are supposedly formal features of discursive judgment as such, as themselves parts of the "content" of judgment. Nevertheless, this becomes understandable once one sees that Kant's point is that modality, unlike the other function types, has no designated role within the judgment itself (as is evidenced by the fact that "All bodies are divisible" is a perfectly well-formed judgment with no reference to modality). Accordingly, one might say that the first three titles are exemplified in the "formal content" of the judgment, that is, its judgmental structure. Moreover, as such, they are expressed in the judgment by modifiers like "all" or "some" and connectives like "if-then," "either-or."

By contrast, modality, which on the Kantian account is "invisible" in the propositional form of the judgment, is said to concern the "value of the copula in relation to thinking in general."<sup>14</sup> This brings out the idea that, rather than referring to an intrinsic feature of a given judgment, viewed as an epistemic unit, modality is concerned with the relation between a particular judgment and a given body of knowledge. It thus concerns what one might term the "epistemic value" of a judgment ("the value of the copula"), and the function of modality is to integrate a judgment within a presupposed system of knowledge. That is why, unlike the other function types, it does not form part of the "content" of a judgment and has no direct propositional expression in it.<sup>15</sup>

## **B. The Completeness of the Moments under Each Title**

The second part of the completeness thesis, namely, that there are precisely three moments falling under each of the four titles, is considerably more controversial. Indeed, William and Martha Kneale undoubtedly speak for many when they remark that "[t]he fact that he [Kant] is able to provide three species under each heading is obviously a matter of accident. For the species in any one set are not really co-ordinate, and there is no common principle requiring trichotomy."<sup>16</sup>

Recent scholarship, however, has raised the possibility of a more positive assessment of Kant's notorious trichotomies, even though it has not yielded a generally accepted solution to the problem. On the one hand, Reinhardt Brandt, while denying that there is anything like a strict proof, argues that the case for the systematic, non-arbitrary nature of the trichotomous arrangement is to be found in the remarks that Kant attached to the table (A71 / B96–A76 / B101).<sup>17</sup> On the other hand, Michael Wolff takes Kant at his word when he suggests that these remarks were intended merely to guard against certain misunderstandings and insists, contra Brandt, that there is a genuine "completeness proof."<sup>18</sup>

In considering this complex issue, we shall begin with Kant's explication of the table. Here particular attention is usually paid to the first two headings, where Kant seems to import considerations from transcendental logic in order to generate his trichotomies. The obvious problem is that any appeal to transcendental logic in order to derive the moments of judgment in general logic would be viciously circular, since the whole purpose of the Metaphysical Deduction is to derive the categories of the former from the logical functions of judgment specified (supposedly on independent grounds) in the latter.

Kant raises the specter of such a circle in his explication of the moments of quantity by noting that logicians rightly claim that in the use of judgments in syllogisms, singular judgments can be treated like universal ones. This suggests that a consideration of quantity merely from the standpoint of general logic would yield the dichotomy: universal and particular, thereby obviating the need to recognize the singular judgment as a logically distinct form. One might well argue for the inclusion of the singular judgment on the grounds that it has a distinct linguistic expression, but Kant builds his case instead on these initially suspicious grounds:

If . . . we compare a singular judgment with a generally valid one, merely as cognition, with respect to quantity, then the former relates to the latter as unity relates to infinity, and is therefore in itself essentially different from the latter. Therefore, if I consider a singular judgment (*judicium singulare*) not only with respect to its internal validity, but also, as cognition in general, with respect to the quantity it has in comparison with other cognitions, then it is surely different from generally valid judgments (*judicia communia*), and deserves a special place in a complete table of the moments of thinking in general (though obviously not in that logic that is limited only to the use of judgments with respect to each other). [A71 / B96–97]

What makes this seem initially suspicious is the reference to "cognition in general" as opposed to "internal validity." Is not the former supposed to be the

domain of transcendental logic and is not general logic, which abstracts from the properly epistemological question of the relation of cognition to its object, limited precisely to the latter?<sup>19</sup> Although the reference to infinity here is mysterious (why not totality?), Kant's main point is relatively clear and does not involve a vicious circularity. First, general logic is not limited to syllogistics. It also includes a doctrine of judgment as such (indeed, this is its core) and is properly concerned with the forms of the latter considered as "functions of unity." Second, the fact that two moments may be treated as equivalent from the standpoint of syllogistics does not mean that there is no distinction to be drawn between them. In short, Kant's position seems to be that, while the inclusion of the singular judgment as a separate moment is important largely because of the distinct epistemic function of such judgments, its distinction from the particular judgment is still one that falls within the domain of general logic.

A similar analysis is applicable to the moments of quality. The basic dichotomy is between affirmative and negative judgments, to which Kant adds the infinite judgment, which is an affirmative judgment with a negative predicate ("The soul is not mortal"). In explaining this addition, Kant remarks that "in a transcendental logic infinite judgments must also be distinguished from affirmative ones, even though in general logic they are rightly included with the latter and do not constitute a special member of the classification" (A72 / B97). Since Kant here refers to general logic as such (rather than merely to syllogistics), the explanation used in the case of quantity is unavailable. Moreover, the circularity worry is further exacerbated by Kant's explanation, which consists in pointing out that infinite judgments perform a limitative function that is quite distinct from both affirmation and negation, and which therefore "must not be omitted from the transcendental table of all moments of thinking in judgments, since the function of understanding that is thereby exercised may perhaps be important to the field of its pure *a priori* cognition" (A73 / B98).<sup>20</sup>

In order to save Kant from the circularity charge here, it is necessary to distinguish between what general logic is primarily concerned with and what it is capable of providing. In other words, the fact that the distinction between an affirmative and an infinite judgment may be of no use to general logic as such does not mean that the distinction cannot be drawn within it. Indeed, that it can be is evidenced by contrasting the infinite judgment, "The soul is not mortal," with the straightforwardly affirmative, "The soul is immortal." Clearly, the soul does not belong to the class of immortal things simply in virtue of being excluded from the sphere of the mortal. Inanimate objects, such as stones, are excluded from the latter as well without thereby being included among the immortals.



The moments of relation seem to raise the opposite of the problem encountered in the first two headings. Whereas the problem there was the seemingly gratuitous inclusion of a judgment form, the present problem concerns an exclusion, namely, that of the copulative judgment (*judicium copulativum*), which Kant's contemporary readers would naturally expect to be included.<sup>21</sup> As the name suggests, a copulative judgment is one in which either two (or more) predicates are affirmed (or denied) of a single subject or a single predicate is affirmed (or denied) of two (or more) subjects (R3089 16: 652).<sup>22</sup> An example of the first type is "God has created all things and rules over them," and of the second type "God and one's neighbor [*die Nächste*] should be loved" (R3088 16: 652). Thus, since Kant was well aware of such judgments, the question is why he omitted them from the moments of relation.<sup>23</sup>

The answer lies in the difference between Kant's understanding of the logical function of relation and the distinction between simple and complex judgments it replaces. The latter is a distinction between propositional forms, which has nothing to do with distinct functions of thought in judgment. In other words, it does not address the question of the distinct sub-tasks performed by the understanding in an act of discursive judgment. In the case of relation, we have seen that this sub-task is to connect the elements combined (be they concepts or judgments) in a relation of condition and conditioned in such a way that the connection first constitutes a judgmental or epistemic unit, that is, a proposition that may be affirmed or denied. In the case of the hypothetical and disjunctive forms, the component judgments are merely taken problematically within the judgment (neither affirmed nor denied) and only the connection between them constitutes the proposition (UE 8: 194n; 289).<sup>24</sup> By contrast, the elements of a copulative judgment are already viewed as complete judgmental units or propositions that may be affirmed or denied independently of their connection in the judgment. Consequently, their combination does not from Kant's point of view constitute a distinct moment of thought.<sup>25</sup>

Since the case of the moments of modality is relatively non-controversial, it does not require detailed discussion. The main worry is why Kant does not include impossibility as a distinct modality.<sup>26</sup> The answer lies in his view that impossibility is simply the negation of possibility and not a distinct modality. That he viewed the matter in this way is clear from the corresponding table of modal categories, where he places the modal pair: possibility-impossibility (A80 / B106).

Even if all of this is granted, however, it does not amount to anything like a proof (or even a systematic explanation) of Kant's trichotomous divisions.

Thus, if such is to be provided we must look elsewhere, and, as Michael Wolff has suggested, a good place to look is an obscure footnote that Kant appends to the Introduction in the third *Critique*. As Kant there remarks:

It has been thought suspicious that my divisions in pure philosophy almost always turn out to be threefold. But that is in the nature of the matter [*in der Natur der Sache*]. If a division is made *a priori*, then it will be either analytic, in accordance with the principle of contradiction, and then it is always twofold (*quodlibet ens est aut A aut non A*). Or it is synthetic; and if in this case it is to be derived from concepts *a priori* (not as in mathematics, from *a priori* intuition corresponding to the concept), then, in accordance with what is requisite for synthetic unity in general, namely (1) a condition, (2) something conditioned, (3) the concept that arises from the unification of the conditioned with its condition, the division must necessarily be a trichotomy. [KU 5: 197n; 82–83]

As Wolff points out, the note contains a rather unusual and highly abstract exercise in conceptual analysis.<sup>27</sup> What is being analyzed is the concept of an *a priori* division in “pure philosophy,” that is, one based entirely on concepts, without any appeal to intuition. Mathematical divisions, say between the various possible types of triangle or regular polygons, are likewise *a priori*; but they are determined by the forms of a given figure constructible in pure intuition and, therefore, can have varying numbers of members.<sup>28</sup> By contrast, the possibilities for conceptually based divisions are severely limited. Apparently viewing such a division as itself a kind of judgment, Kant reasons that it must be either analytic or synthetic. Since the former is based on the principle of contradiction, it is always dichotomous. Correlatively, by a “synthetic division” is to be understood one that is not analytic, that is, not based on the principle of contradiction. Consequently, it is presumably not dichotomous.

We are thus led to ask on what principle a synthetic division could be based, assuming that it can be neither a pure intuition nor the principle of contradiction yet is somehow possible *a priori* (that is, by means of a consideration of the relation of the concepts involved). The answer suggested by the note is that the division must conform to the specified conditions of synthetic unity, which require a trichotomy.

Nevertheless, this only shifts the question, since we now want to know why a non-analytic *a priori* division must constitute a synthetic unity. Although Kant does not explicitly address this issue in the note, his view seems to be that, as *a priori*, the division must be complete and that this requires that the disjuncts constitute a synthetic unity.<sup>29</sup> In the case of an analytic division, the

completeness is immediately apparent, because the contradictory opposites ( $A$  or non- $A$ ) exhaust the domain. In the case of a non-analytic division, however, where the disjuncts ( $A$  and  $B$ ) are not related as contradictory opposites, this does not occur. Accordingly, a dichotomy here does not guarantee completeness.

It may be useful here to consider a division from the domain of practical philosophy: that of morally assessable acts. Here the fundamental division is between the required and the prohibited. In Kant's terminology, the former may be considered the "condition" and the latter the "conditioned" (in the sense that it is defined in opposition to the former). But the division is manifestly incomplete, since it leaves out a class of acts that are neither required nor prohibited, namely, the permitted. Moreover, rather than being simply an additional alternative, standing in no discernible relation to the first two, the latter is related positively to each of them, without being reducible to either. It includes the first disjunct under itself, since required acts are clearly permissible; while it shares with the second the property of falling within the class of non-required acts (the contrary of the first disjunct). Moreover, this enables it to serve as a "mediator," thereby producing synthetic unity and completing the division.

Admittedly, the most that this shows is that trichotomies of this form are *sufficient* to produce the required synthetic unity; not that they are also *necessary* and, therefore, that all synthetic divisions in pure philosophy must be trichotomous. Nevertheless, at this juncture Kant would presumably be prepared to argue (as he generally does) by elimination. The crucial point is that, since the usual ground of *a priori* synthesis, namely, pure intuition, has been ruled out by the nature of the division (as one in "pure philosophy" rather than mathematics), we are left with concepts as the only conceivable means to produce synthetic unity. But it seems clear that this task cannot be assigned to more than one concept, since the question of synthetic unity would break out again with regard to these new concepts. Thus, given the terms of his analysis, Kant seems entitled to conclude that a synthetic unity is possible in *such* divisions *only if* the disjuncts opposed as condition and conditioned can be united in a third concept.

It remains to be seen whether this abstract analysis, which Kant introduced in an effort to justify a quite different set of trichotomies, is applicable to the division of logical functions. This requires showing that the three functions falling under each heading constitute a synthetic unity in the designated sense. Although this may seem to be a daunting task, it can be accomplished fairly expeditiously on the basis of what we have already learned.

*Quantity.* As we have seen, universal and particular constitute the initial disjuncts, and the problem concerns the addition of the singular judgment as a distinct “moment of thought,” which yet somehow combines these disjuncts. The commonality between the singular and the universal judgment may be seen from their functional equivalence within syllogistics, while that of the singular with the particular consists in their joint occupation of the domain of the non-universal. Thus, the function of singularity stands to the functions of universality and particularity in a manner analogous to that of the morally permissible to the required and the prohibited. Consequently, it completes the division of the moments of quantity by enabling them to constitute a synthetic unity.

*Quality.* Here the initial dichotomy is between affirmation and negation (condition and conditioned), and the infinite judgment is assigned a distinct limitative function. Thus, what must be shown is that the infinite judgment combines the first two, enabling the division to constitute a synthetic unity. And once again, it is clear from the previous consideration of this function that it does. As we have seen, it shares with negation the function of excluding a predicate from the domain of the subject, but it does so through an affirmation regarding the subject. Accordingly, here again the conditions of synthetic unity are met.

*Relation.* Here the situation is more complex, since the initial problem was not so much with the inclusion of the disjunctive function as with the exclusion of the copulative one. Nevertheless, it can be maintained that the conditions of synthetic unity apply here as well.<sup>30</sup> The basic opposition is between the categorical and the hypothetical functions. In the former something is affirmed (or denied) unconditionally, and in the latter only under a condition. But the disjunctive function shares with the categorical the function of affirming (or denying) something unconditionally—to cite Kant’s example, “The world exists either through blind chance, or through inner necessity, or through an external cause” (A73 / B93)—while sharing with the hypothetical the function of making its claim on the basis of a relation of propositions, which, taken individually, are merely problematic.

*Modality.* The primary opposition is between merely problematic and assertoric judgments. The former (which serve as components of hypothetical and disjunctive judgments) are deemed capable of a truth value, though, as prob-

lematic, this value is undetermined. By contrast, the truth value of the latter is determined by their connection with principles of the understanding. Apodictic judgments, however, are those whose truth is determined simply on the basis of their capacity to be true, which is what gives them their status as necessary truths. As Kant puts the matter in his discussion of the modal categories, “**Necessity** is nothing other than the existence that is given by possibility itself” (BIII). Thus, here again, the third function combines essential features of the other two, thereby producing a synthetic unity.

## II. FORMS AND FUNCTIONS OF JUDGMENT

Of itself, however, the demonstration of the completeness of the table of judgments does not enable us to understand the connection between the forms of judgment contained therein and the categories. Moreover, we have seen that critics like Strawson, appealing to the conception of logical form operative in modern logic, question not merely Kant’s catalogue of these forms but also, and primarily, the whole project of moving from them to anything like Kantian categories.

The response to this line of objection, which is also intended as an account of Kant’s procedure, falls into two parts. First, it must be insisted that the modern conception of logical form cannot be seen simply as a replacement for the Kantian conception and, therefore, cannot be appealed to in order to undermine the feasibility of Kant’s own project. Second, we shall see that the Strawsonian criticism is based on a serious misunderstanding of the relationship between logical forms (in Kant’s sense) and the categories.

With regard to the first point, we saw in chapter 4 that Kant understands by “forms of judgment” the basic forms or modes of discursive thinking. As Béatrice Longuenesse points out, this reflects Kant’s rather traditional conception of the subject matter of logic as the universal rules of discursive thinking.<sup>31</sup> She further emphasizes that Kant’s primary focus is on the forms of *judging* as a mental activity rather than on the forms of the *judgments* resulting from this activity.<sup>32</sup> Although the latter provides the starting point of Kant’s account in the “Clue,” it is really directed to uncovering the former. Moreover, this enables us to appreciate the radical gulf separating Kant’s conception of “logical form” from the one operative in modern logic, where, as Longuenesse notes, it generally refers to the logical constants and the rules for their derivation and combination assumed in a given logical calculus.<sup>33</sup>

Again following Longuenesse, we can see that this gulf also enables us to appreciate the irrelevance of a critique of Kant's procedure based on the latter conception of logical form. Although Strawson is undoubtedly correct in denying that *this* notion of logical form can plausibly be taken to provide a clue to the discovery of a privileged set of concepts, it hardly follows that the Kantian conception, properly understood, cannot provide one. Nor may it be argued, as it often is, that the great formalizing power of modern logic of itself totally invalidates the Kantian approach. This great power is not to be denied; the point is rather that truth-functional logic and quantification theory have little to add to the analysis of discursive thinking *per se*. Moreover, this is no accident, since logicians from Frege to the present tend to dismiss any such concern as psychological and, therefore, as not constituting part of logic properly construed.<sup>34</sup>

Unfashionable as it may be, however, it is a fundamental premise of this book that Kant's concern with mental acts is not to be construed in a psychological sense. Or, if one insists that *any* account of mental acts is by definition psychological, then the claim is that the account is not psychological in a pejorative sense. By the latter is meant not only one that views Kant as engaged in some illicit form of metaphysical reflection about a noumenal self and its super-empirical activities (Strawson's "imaginary subject of transcendental psychology"),<sup>35</sup> but also one that sees Kant as offering a naturalized, empirical cognitive psychology, which undermines the essentially normative nature of his account of mental activity.<sup>36</sup> As transcendental philosopher, Kant's concern is with the conditions of discursive cognition as such. Consequently, if, as Kant maintains (and Strawson apparently concurs), human cognition is discursive, it seems at the very least plausible to look to the nature of judgment in order to uncover the intellectual conditions of such cognition.

Assuming this, together with the previously noted distinction between forms of judging and forms of judgment, the suggestion is that the latter provide a clue to the understanding of the former and these, in turn, are the key to discovering the categories. Essential to this account is the relation between the notions of "form" and "function" of judgment. Although intimately related (indeed, Kant often uses these terms interchangeably), there is, as Longuenesse points out, an important distinction to be drawn between them, which is essentially one of process or activity and product.<sup>37</sup> In other words, the specific "forms" of thought or judgment arise from the various expressions of the generic "function" of thinking or judging, which Kant describes as "the unity of the action of bringing different representations under a common one" (A68 /

B93). Thus, with due apology to Frank Lloyd Wright, one might say that for Kant “form follows function.”

In order to see the relevance of this to our present concerns, it is important to recall Kant’s definition of ‘function’ as “the unity of the action of ordering different representations under a common one” (A68 / B93). The “action” in question is judgment, and its “unity” is the underlying rule in accordance with which the different representations are connected in a judgment.<sup>38</sup> Since this rule is itself a concept or way of conceptualizing, it suggests the possibility that there may be such a rule embedded in every judgmental form, specifying the manner in which the representations must be connected insofar as one judges under that form. Such a concept rule would not be a category, since it is not a “concept of an object in general”; but it might appropriately be characterized as “precatoreal” or even “proleptically catoreal,” in much the same sense as we saw in chapter 4 that an unconceptualized intuition might be viewed as proleptically a representation of an individual.

Although a similar analysis can be extended to all twelve of Kant’s judgment forms, we shall limit our consideration to the three relational ones: the categorical, hypothetical, and disjunctive forms. Since Kant explicitly connects these with the three most important categories (substance, causality, and reciprocity or community), an analysis of the way in which these forms “follow” or embody conceptual functions should suffice for present purposes.

To begin with the least controversial, it seems clear that the exercise of the categorical function requires the concept of a subject of which properties may be either affirmed or denied and, therefore, a capacity to distinguish between a subject and its properties. Correlatively, the subject of a categorical judgment (the object judged about) is always conceived of as a bearer of properties. For example, in the judgment “Socrates is mortal,” the subject (Socrates) is conceived of as the owner of a property (mortality), which requires the distinction between a subject and its properties. Unless one could do this, one could not form the judgment.

It does not follow from this, however, that one must possess or apply the pure concept of substance, which, in one place, Kant defines as the concept of something that can be conceived of only as subject, and never as predicate of something else (B129),<sup>39</sup> in order to make judgments of the categorical form. This is not merely the concept of something that *can* serve as the bearer of properties but of something that *must* always be conceived of in that capacity, that is, of something that must always be taken substantively. But it is obvi-

ously not the case that this concept is required in order to judge categorically, since we can make perfectly good categorical judgments about properties and abstract objects, as well as the usual substance candidates.<sup>40</sup>

Nevertheless, this does not have the negative implications that Kant's critics tend to assume. The logical function is a rule for the conceptualization of the content of a categorical judgment. It expresses the necessity of conceiving of the subject of such a judgment as a bearer of properties and, therefore, as not itself a property of something else. This amounts to the claim that, in order to judge categorically, it is necessary to consider the subject *as if* it were a substance, not, to be sure, in the full-blown ontological sense but in the logical sense that within the judgment it must be taken substantively. The rule "never a predicate of anything else" thus applies *within* a given judgment.

By contrast, the ontological concept or category of substance may be characterized as the thought of some entity that must be conceived of as subject in *every* judgmental context. Although it certainly follows from this that the logical function or "judgmental concept" operative in the act of making judgments of the categorical form is not to be equated with the ontological concept or category, it also follows that the two are intimately related. Indeed, if this analysis is correct, the latter arises from the former through a kind of hypostatization or projection onto an object, which may or may not be legitimate in a particular instance.<sup>41</sup>

The analysis of the hypothetical form of judgment requires a somewhat different treatment, but it yields a similar conclusion. To begin with, this form should not be construed in truth-functional terms as the material conditional.<sup>42</sup> To judge hypothetically, in the sense in which such a judgment is relevant to Kant's argument, is to assert a connection between two propositions, such that the assumption of the truth of the one justifies the inference to the other. Each of these propositions, taken individually, is viewed as merely problematic; the judgment asserts only the conceptual connection between them. Kant expresses this by remarking, "It is only the implication [*Konsequenz*] that is thought by means of this judgment" (A73 / B98).

It follows from this that the exercise of the hypothetical function presupposes what one might term an "ordering rule" for the sequence of propositions linked in the judgment. Moreover, this rule may be defined as the relation of ground and consequent. Just as it is necessary in the case of the categorical judgment to determine which element is to be thought of as subject and which as predicate, so it is necessary in the case of the hypothetical judgment to deter-



mine which of the propositions is to provide the ground or basis ("inference ticket," in Ryle's sense) for the assertion (or negation) of the other.<sup>43</sup> Accordingly, the claim is that the concept of the relation of ground and consequent is the rule embedded in judgments of the hypothetical form. The justification for this is that to judge hypothetically just is to link the items connected in the judgment in accordance with this rule.

The point is nicely illustrated by Kant's own example of a hypothetical judgment: "If there is perfect justice, then obstinate evil will be punished" (A73/B98). First, this judgment expresses a connection between the thoughts of two states of affairs, neither of which is deemed in the judgment to be actual. As such, it fits the previous characterization of the hypothetical form. Second, these two problematically conceived states of affairs are thought as connected in such a way that the assertion of the first provides a ground for the assertion of the second. In short, the judgment expresses the thought of the dependence of a state of affairs in which the obstinately wicked are punished on one in which there is perfect justice.

This dependence need not, however, be understood in causal terms. The reason we assume that the obstinately wicked will be punished in a world in which there is perfect justice is not that we presuppose that the latter must contain some causal mechanism capable of accomplishing this task (although it might), but simply that punishment for the obstinately wicked constitutes part of the concept or description of a world in which there is perfect justice (at least it constitutes part of Kant's concept of such a world). In short, the judgment is analytic; the concept of such a world provides the logical (though not the "real") ground of the punishment of the obstinately wicked.

This shows that we can dismiss as misguided Guyer's worry when he remarks, "[I]t is hard to see why we should be able to make hypothetical—that is, 'if. . . then ——'—judgments only if we can detect causal connections among objects. . . ."<sup>44</sup> Like similar remarks in Strawson and other dismissive critics, this is correct but beside the point. For contrary to this widely shared assumption, Kant's claim is not that we can judge hypothetically about the world only insofar as we possess and apply the concept of causality; it is rather that the exercise of the hypothetical function necessarily involves (indeed, consists in) an ordering of problematic judgments by means of the relation of ground and consequent.

This does not, however, undermine the move from judgmental form to pure concept, which is the concern of the *Metaphysical Deduction* as a whole. It

shows instead that this move is indirect, mediated by the analysis of judgment in which to judge under a given form is to exercise a certain logical function. Although the logical relation of ground and consequent is not equivalent to the causal relation (this was the mistake of rationalists, such as Spinoza, as well as occasionalists who limited genuine causality to God), it is arguably a necessary condition of the possession of the latter concept. Since this concept is just that of the relation of a *real* (rather than a merely) logical ground to its consequent, one could not have the concept in this form unless one already had the generic concept of the relation. Moreover, this closely parallels the connection between the categorical form of judgment and the pure concept of substance. In both cases, the move from judgmental form to category turns on an ontological application of the logical function embedded in all judgments of the corresponding form.

Unfortunately, things are not as straightforward with regard to the correlation between the disjunctive function and the pure concept of community. Indeed, Kant himself recognizes that the correlation in this case is far from obvious, and in the second edition he adds an explicit defense of it (BIII–13). As it stands, however, this defense is not entirely convincing because of an unclarity concerning the nature of the logical function involved. Kant begins by noting that in a disjunctive judgment the elements combined (problematic judgments) are viewed as constituting a whole (in the sense that they exhaust the possibilities), and he suggests an analogy between this and the thought of a collection of things as constituting a whole. But this analogy breaks down because in the case of a disjunctive judgment, which Kant understands as an exclusive disjunction, the assertion of one element entails the negation of the others, whereas in the case of the pure concept, which involves the thought of reciprocal connection, the assertion of one element entails the assertion of the others. In short, the logical function appears to be that of exclusion, whereas the corresponding category is the concept of a reciprocal connection.

Nevertheless, a connection between them may be preserved, if we take the logical function involved in disjunctive judgment to be coordination rather than exclusion. Moreover, there is support for this reading in the fact that Kant contrasts the relation of coordination expressed in judging under this form with that of the subordination operative in hypothetical judgments. As Kant suggests, this is analogous to the (ontological) coordination of items thought under the category of community. Accordingly, the case for a correspondence between judgment form and category can be made here as well, provided that we focus on the underlying logical function.<sup>45</sup>

### III. THE METAPHYSICAL DEDUCTION "PROPER": FROM GENERAL TO TRANSCENDENTAL LOGIC

The preceding analysis fell within the domain of general logic, though in discussing the connection between logical function and category I oriented it toward transcendental logic. The aim of what is here termed the metaphysical deduction proper is to ground the connection between the two "logics" on the basis of a principle, thereby justifying the claim that the judgmental forms of the former provide the "clue" to the discovery of the basic concepts of the latter.

The underlying principle is that it is one and the same understanding, governed by the same set of rules or functions, that is operative in both domains. In other words, the argument turns on an assumed isomorphism between general and transcendental logic and the functions of thought analyzed in each. Or, more precisely, it turns on an isomorphism between the logical and the real use of the understanding, that is, between its use in judgment, in which it connects pre-given representations by bringing them under concepts, and its use in determining sensible intuition, thereby generating a determinate content of thought.<sup>46</sup>

This strategy involves two closely related and significant drawbacks. One is an unwelcome additional level of obscurity. In introducing the radically new thesis that the understanding has a real use (as a condition of cognition), Kant is forced to appeal in the first five paragraphs of §10 to some of the central and most difficult aspects of his transcendental account of cognition, including the doctrine of synthesis and the respective transcendental functions of the imagination and the understanding. The former is said to synthesize our representations, and the latter to "bring this synthesis to concepts" (A78/B103). Since Kant does little to prepare the reader for any of this, much of what he says in these paragraphs seems like a series of bald assertions, which are only intelligible in the light of his subsequent discussion in the Transcendental Deduction.

The second problem concerns the apparent circularity of this approach. Even if the preliminary discussions of the transcendental functions of the imagination and understanding were sufficiently intelligible in their own terms, the fact remains that Kant has not yet shown that these faculties have a real or transcendental use or, more generally, that there is a subject matter for a transcendental logic. Once again, Kant only purports to establish this in the Transcendental Deduction. Consequently, the Metaphysical Deduction presupposes the results of the Transcendental Deduction, while the latter, in turn, presupposes the results of the former, since it begins with the assumption that a defini-

tive set of pure concepts or categories has already been established as a matter of fact (the *quid facti*) and proceeds to address the question of their validity (the *quid juris*).<sup>47</sup>

Nevertheless, neither of these problems creates an insuperable difficulty. First, even though the obscurity of many of the details of Kant's account in the first five paragraphs of §10 is undeniable, it does not preclude an appreciation of the basic point he is trying to make in the section as a whole. Second, the circularity is not vicious, since Kant's account of the real or transcendental function of the understanding, which is to be contrasted with its previously explicated logical function in judging, is clearly intended as anticipatory or provisional in nature.<sup>48</sup> In other words, we may take Kant to be arguing conditionally that *if* the understanding has a real or transcendental function, that is, if there is such a thing as a transcendental logic (something only to be established subsequently), then we are entitled to assume that its pure concepts or categories will correspond to the logical functions operative in the activity of judgment as analyzed in general logic. The nerve of this argument is contained in the crucial sixth paragraph, where Kant writes:



The same function that gives unity to the different representations **in a judgment** also gives unity to the mere synthesis of different representations **in an intuition**, which, expressed generally, is called the pure concept of the understanding. The same understanding, therefore, and indeed by means of the very same actions through which it brings the logical form of a judgment into concepts by means of the analytical unity, also brings a transcendental content into its representations by means of the synthetic unity of the manifold in intuition in general, on account of which they are called pure concepts of the understanding that pertain to objects *a priori*; this can never be accomplished by universal logic. [A79 / B104–105]



Although the contrary has frequently been thought to be the case, the central claim of this paragraph is the identity of the understanding and its activity (function) as considered in general and in transcendental logic.<sup>49</sup> Thus, the first sentence speaks unambiguously of the “same function” producing unity in both judgment and intuition, and the second sentence refers to the “same understanding” as well as to the “same operations” thereof. We have already considered at some length the logical operation of the understanding and have seen that it involves the unification of representations under concepts. We have also seen that this unification takes place in certain determinate ways, which can be called “forms” or “functions” of unity. Once again, assuming that the understanding has a real as well as a logical use, what Kant is now claiming is that this



same unifying function also takes place at the level of intuition, thereby providing the representational content presupposed by the understanding in its logical activity.

One of the factors that has frequently misled commentators here is Kant's contrast between analytic and synthetic unity, which is sometimes taken as indicating that he is contrasting the activity of the understanding in forming analytic judgments, supposedly the concern of general logic, with its activity in synthetic judgments, which is the business of transcendental logic.<sup>50</sup> But such a reading is not warranted by the text. First, "analytic unity" refers to the concepts that are united in judgments. Nowhere does Kant maintain that judgments themselves are analytic unities.<sup>51</sup> Second, as we have already seen, concepts are analytic unities because they unite in a single representation a series of marks that pertain to a diversity of objects. In fact, it is precisely because concepts are such unities that they can be combined with one another in judgments, whereby "many possible cognitions are . . . drawn into one" (A69 / B94). Once again, this holds true whether the judgment is analytic or synthetic. Third, by the "logical form of a judgment" Kant means a judgment of a given logical form. Thus, Kant's cryptic claim that by means of analytic unity the understanding "brings the logical form of a judgment into concepts" means simply that the understanding produces a judgment of a specific logical form by combining its concepts (analytic unities) in a determinate manner. Insofar as the understanding produces judgments, or judges, it also produces the forms of judgment.<sup>52</sup> The table of logical functions supposedly contains the complete specification of these forms.

Kant further contends that the "same understanding" also "introduces a transcendental content into its representations by means of the synthetic unity of the manifold in intuition in general." The expression "transcendental content" is obscure and subject to a variety of interpretations. Nevertheless, the most plausible is to take it as referring to the just-mentioned synthetic unity of the manifold.<sup>53</sup> This is so even though the text states that the transcendental content is introduced *by means of* the synthetic unity, not that it *is* this unity. Strictly speaking, a transcendental content is an extra-logical, objective content, that is, one that involves relation to an object. The key point here, which Kant only develops in the Transcendental Deduction, is that the synthetic unity of the manifold brought about by the categories is the form of the thought of an object in general. Consequently, insofar as the understanding produces such a synthetic unity, it also relates its representations to an object, thereby introducing a transcendental content. Since the determination of this

synthetic unity is isomorphic with the discursive act of judgment, this enables Kant to talk about the “same operations” or, more generally, to present a picture of the understanding as engaged in one fundamental activity of unification occurring at two levels. Finally, the reference to “intuition in general” is intended to indicate that this general transcendental or objectifying function of the understanding is independent of the particular nature of the manifold of intuition.


Assuming, then, that the understanding has such a function, and that it exercises it through the same operations through which it judges, it follows that the logical functions of judgment, which are the forms in accordance with which the understanding unites its concepts in judgment, will also be the forms in accordance with which it unites the manifold of intuition in order to determine an object for judgment. In short, the pure concepts of the understanding, which introduce the requisite transcendental content, are nothing other than the logical functions of judgment, viewed in connection with the manifold of intuition. Moreover, this allows Kant to conclude in the final step of the Metaphysical Deduction,




In such a way there arise exactly as many pure concepts of the understanding, which apply to objects of intuition in general *a priori*, as there were logical functions of all possible judgments in the previous table: for the understanding is completely exhausted and its capacity entirely measured by these functions. Following Aristotle we will call these concepts **categories**, for our aim is basically identical with his although very distant from it in execution. [A79–80 / B105]

This “deduction” turns on the quasi-identification of the logical functions and the pure concepts. Rather than constituting two distinct sets of concepts—one pertaining to judgment, the other to sensible intuition—that must somehow be brought together, there is a single set of functions belonging to a single understanding operating at two levels. Although Kant does not express himself in precisely this way in the first edition of the *Critique*, he does so often enough elsewhere to remove any doubts that this is his view. For example, in the B-Deduction he writes: “But now the categories are nothing other than these very functions for judging, insofar as the manifold of a given intuition is determined with regard to them” (B143).<sup>54</sup>

This claim, which is central to the understanding of both the Metaphysical and Transcendental Deductions, underscores the significance of the logical functions in Kant’s account of cognition, which is just what we would expect, given the discursivity thesis and the status of these functions as fundamental

forms of discursive thought. It also confirms the analysis offered in the second part of this chapter, according to which the logical functions, regarded as forms of conceptualization embedded in the various judgment forms, are characterized as “precatagoreal.” As we can now see more clearly, they are such because they express at the level of judgment the same function of thought that the category expresses at the level of intuition. This is precisely what makes them indispensable as “clues” to the discovery of the latter. Contrary to the view of Strawson, Guyer, and other critics of the *Metaphysical Deduction*, it is not that we need the category in order to be able to judge under a certain form; it is rather that we can be in possession of a given category only because we are capable of judging under the corresponding form. 

Nevertheless, it is equally important to keep in mind the distinction between the logical functions and the categories, which is why their relation has been characterized as one of quasi-identity. The distinction is, however, functional rather than substantive. As Kant makes clear, the categories are to be equated not simply with the logical functions but with these functions qua operating at the level of intuition and introducing a “transcendental content” into the manifold of intuition. Thus, a reference to sensible intuition (though not a particular type thereof) is an essential component of the very concept of a category for Kant, whereas it is completely alien to the concept of a logical function. This reference to sensible intuition will eventually lead to the need to find schemata in our sensible intuition for these categories. But a prior and more basic problem is the previously noted need to show that the forms of discursive thought, which find their judgmental expression in the logical functions, also have an objectivating role to play in connection with sensible intuition. This is the task of the *Transcendental Deduction* to which we now turn. 


## **Part Three Categories, Schemata, and Experience**




## Chapter 7 The Transcendental Deduction


In the Preface to the first edition of the *Critique* Kant remarks that the set of investigations contained in the Transcendental Deduction cost him more labor than any other (Axvi). He then proceeds to add to these labors, as well as to those of his commentators, by completely recasting the argument in the second edition. Partly as a labor-saving device and partly for philosophical reasons, I shall here focus on the later version. The justification for this lies in the fact that the B-Deduction argument is structured in such a way as to make it evident that the central problem is the demonstration of a connection between the intellectual and sensible conditions of human cognition. Although this is likewise true of the A-Deduction, it is obscured there by the way in which Kant presents his argument.<sup>1</sup> Thus, by concentrating on the later version we can consider Kant's solution to the problem in its most perspicuous form.

The problem itself is clearly stated in a passage from the introductory portion of the Deduction that is contained in both editions. It consists in the worry that, for all that has been shown so far,

appearances could . . . be so constituted that the understanding would not find them in accord with the conditions of its unity, and everything would then lie in such confusion that, e.g., in the succession of appearances nothing would offer itself that would furnish a rule of synthesis and thus correspond to the concept of cause and effect, so that this concept would therefore be entirely empty, nugatory, and without significance. Appearances would nonetheless offer objects to our intuition, for intuition by no means requires the functions of thinking. [A90–91 / B123] 

The possibility to which Kant here alludes calls to mind Descartes's notorious specter of a malignant genius, who systematically deceives us regarding our most evident cognitions. Since the central task of epistemology for more than three centuries has been to answer the global, external world skepticism implicit in this specter and its modern variants, it is frequently assumed that Kant is likewise engaged in such a project in the Deduction. Moreover, judged by this standard the Deduction is generally deemed a failure. 

This reading, however, involves a conflation of the project of the Deduction with that of the Refutation of Idealism. Kant's worry in the former is analogous to the Cartesian one, in that both are concerned with what might be termed a "cognitive fit." Nevertheless, they differ radically in their understanding of the ingredients in this fit.<sup>2</sup> On the Cartesian view, these are thought and being, an inner, subjective realm of cognitions in which certainty is attainable, and an elusive external and objective world. Thus, the worry concerns the correspondence between our evident cognitions and an *an sich* reality. By contrast, for Kant the ingredients are two species of representation, and the worry is that the deliverances of sensibility might not correspond to the *a priori* rules of thought. Accordingly, the Kantian specter is one of cognitive emptiness rather than of global skepticism.

This clarification of the underlying problem enables us to deal with the basic exegetical difficulty concerning the B-Deduction: the division of the argument into two parts, each of which presumably establishes the necessity of the categories. The first part (§15–§21) asserts their necessity with respect to objects of sensible intuition in general. The claim is that any sensible content must be subject to the categories if it is to be brought to the unity of consciousness. The second part (§22–§27) argues for the necessity of the categories with respect to human sensibility and its objects. 

By emphasizing the systematic importance of this structural feature of Kant's argument, Dieter Henrich has set the agenda for all subsequent work on the B-Deduction. Arguing against older commentators, who endeavored to interpret it in the light of models taken from the A-Deduction, Henrich insists on the

two-step nature of the proof. Indeed, he stipulates this as a criterion for the success of any interpretation of the B-Deduction:

The interpretation must show that, contrary to the initial impression that the two conclusions merely define the same proposition, . . . sections 20 and 26 offer two arguments with significantly different results, and that these together yield a single proof of the transcendental deduction. We shall call this task the problem of the two-steps-in-one-proof.<sup>3</sup>

Although Henrich's criterion has been widely accepted, his analysis of the argument has not. Recognizing that the overall goal of the Deduction is to eliminate the above-mentioned specter, Henrich views the argument as proceeding in two steps. The first achieves this result for a range of sensible intuitions, namely, those that already possess unity; while the second removes this restriction, thereby eliminating the possibility of non-conformity for all human sensible intuition on the grounds that space and time are unities (as shown in the Transcendental Aesthetic) and that all our intuitions are spatiotemporal.<sup>4</sup>

The criticism of Henrich's interpretation centers on two points. One is his apparent attribution to Kant of the view that intuitions already possess unity, independently of the activity of the understanding.<sup>5</sup> In response, Henrich has clarified his position, stating that he never intended to attribute such a view to Kant. Referring to B144n., where Kant discusses the ground of proof of the first part of the Deduction, Henrich suggests that what is shown in this part is merely that given intuitions contain a unity, insofar as they are related to apperception, but that this relation cannot, without further argument, be affirmed of everything given in sensibility.<sup>6</sup> This is an important clarification, but it leaves in place the assumption that the first part of the B-Deduction affirms the validity of the categories under a restricting condition that is then removed in the second part.<sup>7</sup>

This is the other major point of contention, since it seems counter-intuitive to suggest that the first part of the B-Deduction contains a restriction on the range of application of the categories that is removed in the second.<sup>8</sup> Kant clearly states that the first part is concerned with the relation between the categories and the manifold of sensible intuition in general and the second with their relation to objects of human sensible intuition. But since human sensibility is a species of sensibility in general, it seems more natural to assume that the first part contains the broader rather than the narrower claim.<sup>9</sup>

As has long been recognized, however, on this reading it is difficult to see how the second part of the Deduction provides anything more than a trivial in-



ference from genus to species. If it has been shown that the categories apply to sensible intuition in general, then surely it follows that they must likewise apply to the specifically human form of intuition.<sup>10</sup> Accordingly, the interpretive problem is somewhat more complex than Henrich's analysis suggests. Not only is it necessary to explain how the two steps constitute a single proof, but this must be done in a way that avoids taking the second as a trivial consequence of the first.

The present interpretation is motivated primarily by the attempt to meet these objectives. It turns largely on a sharp distinction between the epistemic functions assigned to the categories in the two parts of the Deduction. Their function in the first part is to serve as rules for the *thought* of an object of sensible intuition in general, that is, as discursive rules for judgment. That is why the argument abstracts from the particular nature of human sensibility and refers to sensible intuition in general. It shows that any representation that is brought to the "objective unity of apperception" is also thereby related to an object in a judgment and, as such, necessarily stands under the categories.

By contrast, the aim of the second part of the Deduction is to establish the applicability of the categories to whatever is given under the conditions of human sensibility. It attempts to do this by demonstrating (through their connection with the imagination) that the categories also have a non-discursive function as conditions under which whatever is given (in accordance with the forms of sensibility) can enter empirical consciousness. In short, it attempts to link the categories (albeit indirectly) to the *perception* rather than merely the *thought* of objects.<sup>11</sup>

The difference from Henrich's reading is illustrated by the different ways in which they address the specter. As already noted, Henrich suggests that the first part of the B-Deduction removes it for a certain range of intuitions (those that already possess unity), while the second part moves beyond this restriction and removes it for all. By contrast, the reading offered here maintains that the first part by itself leaves the specter completely in place, since it is concerned exclusively with the conditions of the *thought* of objects; so this essential task is assigned entirely to the second part.<sup>12</sup>

The chapter is divided into three parts. The first and third are devoted, respectively, to the central arguments of the two steps of the Deduction as characterized above. The second, which serves as a bridge between these two steps, deals with the subjective unity of consciousness. Although initially introduced almost as an afterthought, this conception turns out to be of considerable systematic significance because it, together with the related discussion of judg-

ments of perception in the *Prolegomena*, focuses attention on perceptual, non-discursive modes of consciousness that become the explicit concern in the second part of the Deduction.

## I. APPERCEPTION, SYNTHESIS, AND OBJECTIVITY

### A. The Transcendental Unity of Apperception:

#### The Principle and Its Analyticity

The B-Deduction officially begins with a general account of synthesis or combination, considered as a spontaneous activity of the understanding on materials given from without. Its essential claim is that “we can represent to ourselves [*wir uns . . . vorstellen können*] nothing as combined in the object without having previously combined it ourselves” (B130).<sup>13</sup> For reconstructing the argument, however, the real starting point is Kant’s canonical formulation of the principle of the transcendental unity of apperception. According to this formulation,

[t]he **I think** must **be able** to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible or else at least would be nothing to me. That representation that can be given prior to all thought is called **intuition**. Thus all manifold of intuition has a necessary relation to the **I think** in the same subject in which this manifold is to be encountered. [B131–32]



This constitutes the real starting point of the B-Deduction because it is the first principle on which the first part of the proof is based. As we shall see, the preliminary conclusion Kant draws in §20 must be viewed as a logical consequence of this principle when fully articulated. But even though Kant presents this as a single principle, it embodies at least three distinct claims, of which the last may be seen as a synthesis of the first two.




The first, expressed in the first sentence, applies to each of a subject’s representations taken individually. It asserts that in order for any of these to be anything *to me*, that is, to represent anything *for me*, it must be possible to think it as mine. Since a representation for which this is not possible could not represent anything for me, it would ipso facto be “nothing to me.”


Although Kant is sometimes thought to have denied (fallaciously) that one could have representations or mental states of which one could not be conscious as one’s own and, thus, conflate consciousness and self-consciousness, a

careful examination of the passage and of Kant's overall account of apperception indicate that he does nothing of the sort.<sup>14</sup> First, in spite of the wording of the text, Kant's claim is not that the *I think* must be able to accompany all my representations *tout court*; it is rather that it must be able to do so, *if* they are to function cognitively for me as representations. Moreover, the "I" here is the abstractly conceived cognitive subject, not an empirically real subject with dreams, wishes, desires (both conscious and unconscious), and the like.<sup>15</sup> In fact, Kant is perfectly willing to countenance the possibility of representations (mental contents or states) *in me* that are nothing to *me*, cognitively speaking.<sup>16</sup>

Second, and equally important, Kant is claiming only that the *I think* must *be able to accompany* (my emphasis) all my representations, not that it must actually do so. In other words, Kant is here affirming merely the necessity of the possibility of self-ascription. Moreover, this is clearly true, since a representation that (for some reason) I could not ascribe to myself could not represent anything for me, though it might very well be mine in some non-cognitive sense.<sup>17</sup>

The second prong of the apperception principle applies to a subject's representations taken collectively, insofar as they are thought to constitute a unified thought. Approaching the text in the light of the richer discussion in the A-Deduction, the basic concern is with the representation of a manifold as a manifold.<sup>18</sup> Whatever its content, such a representation has the form of a single complex thought (a synthetic unity) and requires a single thinking subject. The point here is the one noted by William James: a set of distinct thoughts of the elements of a whole can never be equivalent to the thought of the whole itself.<sup>19</sup> Although each of the representations that collectively constitute the single complex thought could conceivably be distributed amongst a multiplicity of thinking subjects, the thought itself could not be. It must be thought by (or ascribed to) a single subject.

As already indicated, the third prong of this principle consists in the unification of the first two. The result of this unification is the principle of the necessary synthetic unity of apperception, which serves as the fundamental premise of the first part of the B-Deduction. Simply put, this principle states that the components of a complex thought must be connected in such a way as to allow for the possibility of their ascription to a single thinking subject, which entails that they constitute a synthetic unity. 

The argument for this principle proceeds by taking the first two prongs in reverse order: since a single complex thought logically requires a single thinker, it follows (1) that each of the components of such a thought must be ascribable to 

an identical thinking subject; and (2) it must (necessarily) be possible for this thinker to be aware of its identity. The latter is a necessary condition of the possibility of a number of discrete representations being united in the thought of a single subject as its representations, and, *a fortiori*, of their constituting a single complex thought. In other words, if representations *A*, *B*, and *C* are to be thought together in a single consciousness, then the I that thinks *A* must be identical to the I that thinks *B*, and so forth. In addition, if the subject is to be conscious of these representations as constituting a unity, then it must also be possible for it to become conscious of its own identity as subject with respect to the thought of each of these representations. As we shall see in more detail shortly, these conditions are reciprocal.



Our immediate concern, however, is with a feature of Kant's account that has been neglected up to this point: the analyticity of the principle of apperception. Kant affirms this twice within the space of four pages: asserting first that, in spite of its significance, "this principle of the necessary unity of apperception is . . . itself identical, thus an analytic proposition" (B135); and referring back to this proposition, he reminds us that it is "itself analytic . . . for it says nothing more than that all **my** representations in any given intuition must stand under the condition under which alone I can ascribe them to the identical self as **my** representations, and thus can grasp them together, as synthetically combined in an apperception, through the general expression **I think**" (B138).

Although many commentators pass over these texts with silence, Guyer is keenly aware of them. Nevertheless, he dismisses the analyticity claim on the grounds that it reflects Kant's confusion about his argumentative strategy, a strategy that is allegedly doomed to failure even when freed from this confusion. In support of this interpretation, Guyer points to a footnote in the A-Deduction, where Kant explicitly affirms that the principle of the unity of apperception is synthetic.<sup>20</sup> As Kant there puts it, "The synthetic proposition, that every different empirical consciousness must be combined into a single self-consciousness is the absolutely first and synthetic principle of our thinking in general" (A117n.). Taking this as equivalent to the principle that Kant characterizes as analytic in the second edition, Guyer charges Kant with a further confusion (or at least a puzzling change of mind) regarding the nature of his own first principle.

In dealing with this issue, we must resist the temptation to save Kant from inconsistency by limiting the scope of the analyticity thesis to a trivial aspect of the principle—say, that my representations must be subject to whatever conditions are necessary in order for them to be mine. As noted above, what Kant ac-

tually claims to be analytic is the complex and apparently contentful proposition that “all my representations in any given intuition must stand under the condition under which alone I can ascribe them to the identical self as my representations, and thus can grasp them together, as synthetically combined in an apperception, through the general expression I think.”

This seems to worsen the problem, however, since it is not clear that this latter claim is even true (not to mention analytic). Can I not ascribe to my identical self thoughts that do not, as such, constitute a synthetic unity: for example, the thought of my present puzzlement regarding Kant’s apperception principle and my ongoing concern over the state of the Red Sox?

The answer is that, in thinking these together as my thoughts, I am thereby bringing them into a synthetic unity. Indeed, I cannot ascribe them to my identical self without *in the very same act* also bringing them into synthetic unity (even if it is only the contingent unity they possess in virtue of being jointly recognized as my thoughts). Consequently, it is a condition of the possibility of the self-ascription of distinct thoughts that they can be brought into a synthetic unity; just as it is a condition of such synthetic unity that the thoughts be ascribable to a single thinking subject. As Kant puts it in a passage in the B-Paralogisms, which is clearly intended to complement the analysis of apperception in the Deduction:

That the I of apperception, consequently in every thought, is a **single thing** [*ein Singular sei*] that cannot be resolved into a plurality of subjects, and hence a logically simple subject, lies already in the concept of thinking [*des Denkens*], and is consequently an analytic proposition: but that does not signify that the thinking I is a simple **substance**, which would be a synthetic proposition. [B407–08]<sup>21</sup>

Here Kant not only insists on the analyticity of the principle but also presents it as lying in the “concept of thinking” (or thought). Applying Kant’s own criteria of analyticity, the concept in which the predicate is “contained” and from which it follows by means of the principle of contradiction is that of thinking or thought. Or, more precisely, since Kant explicitly limits the scope of the apperception principle to a discursive intellect (see B138–39 and 145–46), it follows from the concept of discursive thought.<sup>22</sup> It is because such thinking consists in bringing a manifold of intuition under a concept that it cannot be conceived of apart from a numerically identical subject. Conversely, since the subject is being considered merely as the subject of discursive thought, its identity is inseparable from the synthetic unity of its thought. In short, the



analytic proposition is a claim about how the thinking subject must be thought (or conceive of itself) qua engaged in such activity, and not a metaphysical thesis about the nature of the thing that thinks.

What, then, are we to make of the fact that at A117n, Kant characterizes what seems to be the same principle as synthetic? The short answer is that they are not identical. The synthetic principle states that “every different empirical consciousness must be combined in a single self-consciousness.” Here empirical consciousness is equivalent to perception, and the claim is that distinct perceptions must themselves be linked in a single self-consciousness (identified with “transcendental consciousness” or the “consciousness of myself as original apperception”). Since the latter is deemed necessary for the connection of temporally distinct perceptions and, therefore, for the possibility of experience, it is clearly synthetic. But what makes it synthetic is its connection with time and possible experience, not its connection with synthetic unity as such. By contrast, in the discussion of apperception in the B-Deduction, it is merely the abstract notion of a synthetic unity of the manifold in a given intuition (no matter what its “form” or the manner of its givenness) that is at issue.

The differences in the accounts of apperception reflect Kant’s different procedures in the two versions of the Deduction. The A-Deduction begins with an explicit reference to the temporality of experience. In fact, Kant prefaces his discussion with what he describes as a “general remark on which one must ground everything that follows,” namely,

Wherever our representations may arise . . . as modifications of the mind they nevertheless belong to inner sense, and as such all of our cognitions are in the end subjected to the formal condition of inner sense, namely time, as that in which they must all be ordered, connected, and brought into relations. [A99]

By contrast, the first part of the B-Deduction abstracts from the nature of sensibility, which means that it also abstracts from the very conditions required to ground synthetic judgments. As Kant repeatedly insists, the latter require the relation of concepts to sensible intuitions, and in the present case that requirement cannot be met. Instead of relating thought to human intuition, Kant here relates it to the *concept* of “sensible intuition in general,” which makes the judgment analytic. Thus, given Kant’s account of the conditions of synthetic judgment, he had no choice but to regard the principle elucidated in the first part of the B-Deduction as analytic. Of course, it may not be analytic on a different understanding of analyticity (or syntheticity), but that is beside the point.

## B. Synthesis and the Consciousness of


### Synthesis: Analyticity without Sterility

Nevertheless, a major obstacle to the analytic reading of the apperception principle remains, namely, its connection with the doctrine of synthesis. Whether one views this doctrine with Strawson as part of the “imaginary subject of transcendental psychology,”<sup>23</sup> or with other commentators, who take it as a broadly empirical thesis in cognitive psychology,<sup>24</sup> it seems difficult to regard it as the consequence of an analytical principle. Yet this is precisely what Kant claims when he states that from what has already been established about apperception (“this original combination”), it follows that

this thoroughgoing identity of the apperception of a manifold given in intuition contains a synthesis of the representations, and is possible only through the consciousness of this synthesis. For the empirical consciousness that accompanies different representations is by itself dispersed and without relation to the identity of the subject. The latter relation therefore does not yet come about by my accompanying each representation with consciousness, but rather by my **adding** one representation to the other and being conscious of their synthesis. Therefore it is only because I can combine a manifold of given representations **in one consciousness** that it is possible for me to represent [to myself] the **identity of the consciousness in these representations** itself, i.e., the **analytical** unity of apperception is only possible under the presupposition of some **synthetic** one. [B133–34]<sup>25</sup>

This passage contains two distinct claims, which come together in the conclusion that the analytic unity of apperception presupposes a synthetic unity: (1) that the consciousness of the identity of the *I think* “contains” a synthesis; and (2) that it is possible only through a consciousness of this synthesis. Both, as well as the conclusion, are essential to Kant’s overall argument, and we shall discuss each in turn. In order to do so, however, it is first necessary for us to consider briefly Kant’s initial account of synthesis in §15. The manifold, Kant notes there, can be given in a purely sensible intuition, and the form of this intuition can be regarded as nothing more than the mode in which the subject is affected:

Yet the **combination** [*conjunctio*] of a manifold in general can never come to us through the senses . . . for it is an act of the spontaneity of the power of representation, and, since one must call the latter understanding . . . all combination, whether we are conscious of it or not, whether it is a combination of the manifold of intuition or of several concepts . . . is an action of the understanding, which we would designate with the general title synthesis in order at the same time to draw attention to the


fact that we can represent [to ourselves] nothing as combined in the object without having previously combined it ourselves, and that among all representations combination is the only one that is not given through objects but can be executed only by the subject itself, since it is an act of its self-activity. [B129–30] 

This claim (and its variants) is frequently criticized on the grounds that it rests upon the dubious assumption that what is actually given to the mind is a Humean “mere heap” of impressions or even a “chaos of sensation.”<sup>26</sup> Such criticism is misguided, however, particularly in the context of the B-Deduction, where the operative assumption is merely that we are dealing with a mind for which the manifold must be given, that is, with a discursive rather than an intuitive intellect. As the text makes clear, the problem is to explain how such an intellect can *represent to itself* its data as combined, that is, as constituting a synthetic unity. But this problem arises independently of any assumptions about how the manifold is given. For even if we suppose that the data are already given in an organized or unified fashion, the intellect must still represent to itself or think this “given” unity. Kant puts the point succinctly in a letter to Beck: “We must *compose* [*zusammensetzen*] if we are to represent anything as *composed* [*zusammengesetzt*] (even space and time)” (Br II: 515; 482).

Given these preliminaries, we are now in a position to examine Kant’s two claims about the necessary unity of apperception. Actually, the first has already been considered, in the preceding section, where we saw that a subject cannot think (apperceive) its own identity with respect to distinct representations without in the same act bringing them into a synthetic unity. To think this thought (that of the identity of the *I think*) is to unify the distinct representations in a single consciousness, which is why Kant claims that this thought “contains” a synthesis.


Obviously, the converse does not hold; not every act of unification of representations in a single consciousness produces an actual consciousness of the identity of the *I think*. But this is irrelevant, since all that is required is that the unification allow for the *possibility* of such self-consciousness. Kant terms this self-consciousness “universal” (B132), which is here equivalent to transcendental, because it constitutes what he elsewhere terms the “logical form of all cognition” (A117n.). We shall soon see why.

The claim that apperception is possible only through a *consciousness* of synthesis is more complex, because of the ambiguity that attaches to the term *synthesis*. It can refer either to the act itself or to its product; and Kant apparently intended it in both senses. Understood in the second sense, the claim is un-

problematic. Taking the simplest possible case in which there are only two representations (*A* and *B*), the consciousness of the identity of the *I* that thinks *A* with the *I* that thinks *B* obviously requires an awareness of both *A* and *B* together. As Kant puts it, apart from such consciousness “I would have as multi-colored, diverse a self as I have representations of which I am conscious to myself” (B134). But not only is a single consciousness possible only by means of the combination of these representations; it is itself a consciousness of the product of their combination. At least in this sense, then, apperception (the consciousness of an identical *I think*) necessarily involves the consciousness of the synthesis of representations. 

The claim seems considerably more problematic, however, if consciousness of synthesis is taken to refer to the activity. First, there is the question of its compatibility with what Kant says elsewhere. For example, in the above-cited passage from §15, he refers in passing to “all combination, whether we be conscious of it or not,” which suggests that Kant denies that all such acts necessarily involve a consciousness of the act.<sup>27</sup> Second, taken as a psychological thesis, the claim that the mind must be conscious of its acts of synthesis seems wildly implausible. Accordingly, it is tempting to view this claim as an aberration that requires to be explained away rather than as an essential ingredient in Kant’s transcendental account of cognition.<sup>28</sup>

Nevertheless, this dismissive treatment is unwarranted, since a certain consciousness of the act of synthesis is an ineliminable feature of Kant’s doctrine of apperception.<sup>29</sup> Although it is impossible fully to justify such a sweeping claim at present, the basic points can be expressed succinctly. To begin with, it must be noted that the claim is not unique to the B-Deduction. For example, in a striking passage in the earlier version Kant writes:

[F]or the mind could not possibly think of the identity of itself in the manifoldness of its representations, and indeed think this *a priori*, if it did not have before its eyes [*vor Augen hätte*] the identity of its action [*Handlung*], which subjects all synthesis of apprehension (which is empirical) to a transcendental unity, and first makes possible their connection in accordance with *a priori* rules. [A108] 

In addition to providing a graphic illustration of the presence of this thesis in the A-Deduction, this suggests that the necessity of a consciousness of *acts* of synthesis applies specifically to the operations of the understanding and does not include those of the imagination (here the empirical synthesis of apprehension). Presumably, then, in the passage from §15, where Kant refers to combination (synthesis) of which we might *not* be conscious, he had the latter in

mind. To be sure, Kant there links combination explicitly to the understanding and does not refer to the imagination, but that is because he is focusing on discursive thought as such (what he will later call the “intellectual synthesis”). Moreover, we shall see that when Kant does bring the imagination into the story in the second part of the B-Deduction, he virtually identifies it with the understanding. Given this virtual identification, as well as Kant’s claim in §15 that he is referring to *all* combination, it seems reasonable to assume that he intended it to include the imagination and that it was the latter he had in mind when he referred to a combination of which the mind is *not* conscious.

The A-Deduction passage also agrees with its B-Deduction counterpart in asserting that the consciousness of synthesis is necessary if the mind is to be able to think its own identity in the manifold of its representations. The former passage further suggests that the mind must be able to think its identity *a priori*, that is, in accordance with *a priori* rules (the categories), and this is obviously crucial for the overall argument of the Deduction. But this complication may be ignored for the present, since only the initial claim is essential for understanding the connection between Kant’s doctrine of apperception and the necessity of a consciousness of synthesis. In order to understand this, however, we need to see why the mind’s thought of its own identity with respect to the diversity of its representations is dependent upon, indeed consists in, a consciousness of its unifying activity.



First, by the “identity of its action” (which the mind presumably has “before its eyes”) is to be understood the necessary unity of an act of thought. Just as a single complex thought requires a single thinker, so, too, it requires a unified act of thinking in which all the components of the thought are brought into synthetic unity. Even if I am the same subject that thinks *A* at  $t_1$  and *B* at  $t_2$ , this does not make them contents of a single thought. The latter also requires that I think them together in a single act.

Second, because of the contentlessness of the *I think*, there is literally nothing, apart from the consciousness of the identity of its action (in thinking a complex thought), through which the thinking subject, considered as such, could become aware of its own identity. Expressed schematically, the consciousness of the identity of the I that thinks *A* with the I that thinks *B* can only consist in the consciousness of the identity of its action in thinking together *A* and *B* as its representations. That is why a consciousness of synthesis (considered as activity as well as product) is a necessary condition of apperception, even though the latter requires merely the *possibility* of the self-ascription of one’s representations.

Moreover, Kant regarded this complex claim concerning the connection between apperception and synthesis as equivalent to the thesis that “the **analytical** unity of apperception is only possible under the presupposition of some **synthetic** one” (B133–34). Here, and in the important footnote attached to this claim, Kant begins to forge the link between apperception and the understanding that is central to the Deduction. We have already seen that Kant viewed all general concepts as analytic unities, which means that they contain within a single representation the thought of what is common to a multiplicity of distinct representations. We have also seen that he regarded such concepts as produced by a series of “logical acts,” termed “comparison,” “reflection,” and “abstraction.” What we now must consider is the connection between these claims and the doctrine of apperception.

Two points are highly relevant here. First, this identical *I think*, that is, “the bare representation I,” can be seen as the form or prototype of the analytic unity that pertains to all general concepts. In fact, it *just is* this analytic unity considered in abstraction from all content. Consequently, the *I think* is itself the thought of what is common to all conceptualization, which is what makes it “in all consciousness one and the same” (B132). Second, the act of becoming aware of this identical *I think* is the form of the act of reflection, by means of which the mind grasps the identity in difference in the formation of general concepts. Once again, it is nothing more than the “logical act,” considered in abstraction from all content. The consciousness of this act, that is, the consciousness of synthesis, is, therefore, the consciousness of the form of thinking.<sup>30</sup>

It follows from this that the doctrine of apperception is neither a bit of introspective psychology nor an idealistic ontological thesis concerning the manner in which the mind “creates” the phenomenal world by imposing its forms upon the given sensible data.<sup>31</sup> It is rather a formal model or schema for the analysis of the understanding and its “logical” activities. As such, it specifies the conditions to which all my representations must conform, if they are to function in cognition, that is, to be “something to me,” cognitively speaking. Correlatively, the theory of synthesis implied by this doctrine is an analytical account of the mode of operation of the model. This is precisely what Kant had in mind when, at the end of the above-mentioned footnote, he remarks:

And thus the synthetic unity of apperception is the highest point to which one must affix all use of the understanding, even the whole of logic and, after it, transcendental philosophy; indeed, this faculty is the understanding itself. [B134 n.]

### C. Apperception, Objectivity, and Judgment

Although implicit from the beginning, the identification of apperception (as a faculty for producing synthetic unity) with the understanding is the pivotal move. Given this identification, Kant is in a position to connect apperception with the representation of objects and through this with judgment, which leads, in turn, to its connection with the categories. But as a result of Kant's progressive or synthetic presentation, these implications only emerge gradually. Accordingly, we must continue to follow Kant step by step in his analysis of the *a priori* conditions of discursive cognition.

The first and essential step is to establish a necessary connection (actually a reciprocity) between the synthetic unity of apperception (identified with the understanding) and the representation of objects. The argument for this crucial connection is compressed into a single paragraph:



**Understanding** is, generally speaking, the faculty of **cognitions**. These consist in the determinate relation of given representations to an object. An **object**, however, is that in the concept of which the manifold of a given intuition is **united**. Now, however, all unification of representations requires unity of consciousness in the synthesis of them. Consequently the unity of consciousness is that which alone constitutes the relation of representations to an object, thus their objective validity, and consequently is that which makes them into cognitions and on which even the possibility of the understanding rests. [B137]



As the text indicates, the characteristic activity of the understanding is to relate given representations to an object. Since cognition likewise consists in such a relation, this presumably follows from the definition of the understanding as the faculty of cognition.<sup>32</sup> But this immediately gives rise to the question of what is meant by an object (*Objekt*), and we see that it is defined simply as “that in the concept of which the manifold of a given intuition is united.” This reflects Kant’s “Copernican turn”: first-order talk about objects is replaced by second-order talk about the *concept* of an object and the conditions of the representation of an object. The meaning of ‘object’ is thus determined by an analysis of these conditions. And since the conditions consist in a given manifold of intuition and its unification under a concept, an object may be understood as the correlate of an act of conceptual unification. Thus, whatever is represented through such a synthetic unity counts as an object.

Kant is not here referring to a distinct kind of object. He is concerned rather with *any* object of discursive cognition, qua considered merely from the side of

the understanding. In the language of the A-Deduction, he is concerned with the concept of an "object in general."<sup>33</sup> Just as there is no intuition in general, so there is no object in general; though there are (or at least Kant thought there were) necessary and sufficient conditions for the representation of an object as such.

This conception of an object underlies Kant's attempt to link the unity of consciousness with the representation of objects. The crucial claim is that "the unity of consciousness is that which alone constitutes the relation of representations to an object, thus their objective validity." Since Kant presents this as a direct consequence of the principle that "all unification of representations requires unity of consciousness in the synthesis of them," it might seem that he is guilty of a gross non sequitur, because the latter entails merely that the unity of consciousness is a *necessary* condition for the representation of an object, not that it is also a *sufficient* one. Yet this is precisely what Kant claims when he asserts that this unity is "that which alone constitutes the relations of representation to an object."<sup>34</sup> Indeed, it is what he *must* claim, if he is to establish a necessary connection between the unity of consciousness and the categories.

The appearance of a non sequitur can be avoided, however, if we keep in mind the operative meaning of 'object'. First, it follows from the doctrine of apperception not merely that the synthetic unity of representations under a concept of an object is impossible apart from the unity of apperception but also that the latter is impossible apart from the former. In short, the doctrine affirms a reciprocity between the unity of apperception and the synthetic unity of representations (the unity of consciousness and the consciousness of unity). Second, as the identification of apperception with the understanding indicates, this synthetic unity can be attained only by the unification of representations under the concept of an object. Third, since to unify representations in this way just is to relate them to an object, it follows that relation to an object is likewise a necessary condition of the unity of apperception, which makes the latter a sufficient condition of the former. This allows Kant to claim that

[t]he synthetic unity of consciousness is therefore an objective condition of all cognition, not merely something that I myself need in order to cognize an object but rather something under which every intuition must stand **in order to become an object for me**, since in any other way, and without this synthesis, the manifold would not be united in one consciousness. [B138]

Here Kant emphasizes the epistemic force of the apperception principle. To claim that the synthetic unity of consciousness is "an objective condition of all



cognition" is, in effect, to claim that it is an *objectivating* condition. Kant underscores this point by denying that such unity is "merely something that I myself need in order to cognize an object," which would effectively reduce it to a psychological condition without normative import. Nor should it be assumed that this epistemic force is somehow undermined by the reference in the last phrase to "an object for me." The "me" at issue here is not a particular cognizer but the accusative form of the I of apperception, whose transcendental unity constitutes the form of the thought of an object.

It does not follow from this, however, that the mere de facto union of representations in a single consciousness is sufficient to confer objectivity on this union. In order to dispel any such impression, Kant distinguishes in §18 between an objective and a subjective unity of consciousness. Although the latter notion raises problems of its own, which will occupy us in the next section, we are now concerned merely with the former, defined as "that unity through which all of the manifold given in an intuition is united in a concept of the object" (B139).

As Kant makes clear in §19, an objective unity is produced through an act of judgment. And, by way of contrasting this with a relation of representations in accordance with laws of the reproductive imagination (Humean principles of association), which would constitute a mere subjective unity, he writes:

I find that a judgment is nothing other than the way to bring given cognitions to the objective unity of apperception. That is the aim of the copula *is* in them: to distinguish the objective unity of given representations from the subjective. [B141–42]

When this characterization of judgment was first considered in chapter 4, it was noted that it goes beyond the earlier one Kant sketched in the "Clue" by making objective validity (a normative notion) a definitional feature of judgment rather than a property of some judgments. Although the apparently paradoxical nature of this claim was mitigated considerably by the interpretation of such validity as a capacity for truth or falsity rather than truth, it was impossible at that point to discern the ground of this validity. This is precisely what we learn in the Transcendental Deduction, however, through its connection with the objective unity of apperception.

Kant emphasizes the latter point when he criticizes the explanation of judgment provided by logicians as the "representation of a relation between two concepts" (B140). In addition to applying only to categorical judgments (a comparatively minor defect), its main fault, Kant suggests, lies in its failure to determine wherein this relation consists (B141). His answer is that it consists in

their relation to the objective unity of apperception, an answer that was not available to traditional logicians because it stems from transcendental rather than general logic.

We can see from this that the doctrine of apperception underlies Kant's account of judgment as the fundamental act of discursive thought and provides the basis for its inclusion within the domain of transcendental logic. It is also the case, however, that the analysis of judgment illuminates the preceding account of apperception. In particular, it helps us to understand better Kant's insistence not only on the necessity of a synthesis but also on the consciousness thereof, which seemed so puzzling when initially presented. Looking back on this earlier discussion from the vantage point of the conception of judgment, however, it becomes clear that the synthesis to which Kant there referred was nothing other than the act of judgment. Such an act must be attributed to the spontaneity of the subject, since it consists in a representing to oneself of representations as united in an object. It must also be viewed as norm governed, because it is thought to hold for all consciousness ("consciousness in general" or "universal self-consciousness"), not merely for oneself qua particular consciousness. Moreover, if this is correct, it follows that the "consciousness of synthesis" is the self-consciousness built into the act of thinking.<sup>35</sup>

From the systematic standpoint, however, the significance of the account of judgment in §19 consists in the fact that it provides the basis for the introduction of the categories in §20. Moreover, this constitutes a major improvement over the A-Deduction, which attempted to relate apperception to the categories and the latter to experience without explicitly referring to judgment.<sup>36</sup> Indeed, Kant himself underscored the significance of this point and signaled the direction that the new version of the Deduction would take in a footnote from the Preface to the *Metaphysical Foundations of Natural Science*, where he suggests that the deduction could be accomplished "almost by a single inference from the precisely determined definition of a *judgment* in general (an action through which given representations first become cognitions of an object)" (MAN 4: 475–76n.; 190). Whether we take this as an indication of the fluidity of Kant's views on the Deduction during the period between the two editions of the *Critique* or as an anticipation of the proof strategy he adopted one year later, in the B-Deduction, depends on how much is read into 'almost'. But on either reading it shows that Kant gave to this conception of judgment an essential role in the argument.<sup>37</sup>

As first becomes clear in §20, the reason for this lies in the previously discussed connection between the categories and the logical functions. This brief

section, which constitutes the conclusion of the first part of the B-Deduction, is given the heading: "All sensible intuitions stand under the categories, as conditions under which alone their manifold can come together in one consciousness." The argument consists of the following five steps:

- [1] The manifold that is given in a sensible intuition necessarily belongs under the original synthetic unity of apperception, since through this alone is the **unity** of intuition possible (§17). [2] That action of the understanding, however, through which the manifold of given representations (whether they be intuitions or concepts) is brought under an apperception in general, is the logical function of judgments (§19). [3] Therefore all manifold, insofar as it is given in **one** [*Einer*] empirical intuition, is **determined** in regard to one of the logical functions for judgment, by means of which, namely, it is brought to consciousness in general. [4] But now the **categories** are nothing other than these very functions for judging, insofar as the manifold of a given intuition is determined with regard to them (§13). [5] Thus the manifold in a given intuition also necessarily stands under the categories. [B143]

The first two steps of this argument require little comment, since they simply restate what has supposedly already been shown. The third step is obviously pivotal, though it may be read as spelling out the consequences of the first. So construed, the claim is that insofar as the manifold of a given intuition is grasped as a manifold or brought to the synthetic unity of apperception, it is, *by this very act*, also unified in a judgment. But since the logical functions of judgment are the forms of such unification, it follows that the manifold is also determined with regard to them. In short, for a discursive understanding to think the manifold of a given intuition just is to unify it in a judgment by means of the logical functions. The fourth step trades on the connection between the logical functions and the categories supposedly established in the Metaphysical Deduction. Here the key is that the thought manifold is *determined* (Kant's emphasis), that is, positively ordered (say, as *a-b* rather than as *b-a*) through the judgment. Finally, since the categories just are these logical functions, considered as functioning in this way, it follows that the manifold, insofar as it is thought together in a single consciousness, is necessarily subject to the categories.

This result, which Kant at one point equates with the transcendental deduction (B159), takes us beyond the Metaphysical Deduction by offering grounds to support what the latter was forced to assume, namely, that the understanding has a real (as opposed to a merely logical) use, through which it introduces a "transcendental content" into its representations. In other words, it shows that the understanding has an objectivating function, that it relates given rep-

representations to an object through judgment. Moreover, since this function consists in determining what is given in intuition in a fixed manner for judgment, it also shows that the logical functions, considered as functioning in this way, are categories.

It does not, however, show that appearances, qua given under the forms of space and time, necessarily conform to these categories. Consequently, it does not exorcize the specter that appearances might be "so constituted that the understanding would not find them in accord with the conditions of its unity." As I have already indicated, this is the task of the second part of the B-Deduction, through which the validity of the categories is first secured. But before turning to that, we must still consider the concept of the subjective unity of consciousness and its difference from the objective unity.

## II. THE PROBLEM OF SUBJECTIVE UNITY

We have seen that Kant introduces the distinction between objective and subjective unity in §18 in order to dispel the impression, created by the account of the objectifying function of apperception in §17, that a *de facto* connection of representations in consciousness of itself confers objectivity on these representations. We have also seen that Kant obviates any such worry by indicating that the latter applies only to a certain kind of unity, one that is not passively received but spontaneously produced by the understanding through an activity, which in §19 is identified with judgment.

Although this resolves the immediate problem and tells us what a subjective unity is *not*, namely, one produced through an act of judgment, it does not indicate what it is. Moreover, Kant's cryptic account of such unity is itself highly ambiguous, raising as many questions as it resolves. And the problem is exacerbated when considered in connection with the problematic distinction between judgments of perception and judgments of experience drawn in the *Prolegomena*. Not only do these texts contain significant internal tensions; they appear to be incompatible both with one another and with central tenets of the *Critique*.

In dealing with these related problems, three main claims will be argued for: (1) that there is no contradiction between the *Prolegomena* and the *Critique*, since the latter's distinction between two kinds of unity should be viewed as a change of subject rather than a correction of the former's distinction between two kinds of judgment; (2) that the concept of a subjective unity does not simply correct a false impression created by the analysis of apperception but also

points to a central issue in the second part of the B-Deduction; and (3) that the tensions within each of these texts and the appearance of a conflict between them may be significantly mitigated, if not completely removed, by a consideration of the contexts and methodological constraints of the two accounts.

### A. Judgments of Perception and Judgments of Experience

Kant's discussion of this distinction in the *Prolegomena* is intended as a counterpart to the Transcendental Deduction of the *Critique*. In accordance with the work's analytic method, however, instead of confronting the specter addressed by the Deduction, Kant assumes the actuality of certain synthetic *a priori* judgments (identified with "pure natural science") and investigates the conditions of their possibility. The basic claim is that without appealing to a set of pure concepts, we could not distinguish between how things merely seem to a particular observer under contingent and variable conditions and how they are objectively (intersubjectively) constituted as phenomena. As a means of illustrating this claim, Kant introduces the distinction between the two species of empirical judgment:

*Empirical judgments, insofar as they have objective validity, are judgments of experience; those, however, that are only subjectively valid I call mere judgments of perception. The latter do not require a pure concept of the understanding, but only the logical connection of perceptions in a thinking subject. But the former always demand, in addition to the representations of sensory intuition, special concepts originally generated in the understanding, which are precisely what make the judgment of experience objectively valid.* [Pro 4: 298; 92]



Considered in the light of the B-Deduction, the striking feature of this passage is its suggestion that only one of these two species of empirical judgment possesses the objective validity that according to §19, pertains to judgment as such. Members of the other are merely subjectively valid and do not make use of the categories, though, as well-formed judgments, they involve the logical functions.<sup>38</sup> Kant later suggests that the latter contain a connection of perceptions "in a consciousness of my state," which is contrasted with the connection in "a consciousness in general" that occurs in an objectively valid judgment of experience (Pro 4: 300; 94). An equivalent characterization of this merely subjective mode of unification is that the representations are "related to a consciousness in one subject only" (again in contrast to consciousness in general) (Pro 4: 304; 98).



Kant introduces a further complication by distinguishing between two classes of judgments of perception: those that can and those that cannot become judgments of experience by the application of a pure concept. Examples of the latter include “The room is warm,” “Sugar is sweet,” and “Wormwood is bitter” (Pro 4: 299; 93). Such judgments are inherently subjective because they refer to feeling states or sensations that can never be attributed to the object. As a result, in their case there is no work for a pure concept to do. An example of judgments that can become judgments of experience with the introduction of a pure concept is “If [*Wenn*] the sun shines on the stone it becomes warm.”<sup>39</sup> Kant contends that with the addition of the concept of causality this becomes converted into the objectively valid judgment “The sun *warms* the stone” (Pro 4: 301n.; 95). Kant’s final and most interesting example, “The air is elastic,” is a case in which one and the same proposition is viewed first as expressing a judgment of perception and later a judgment of experience (Pro 4: 299; 94).<sup>40</sup>

Without further pursuing the details of this discussion, we can note that an underlying theme is the contrast between what Kant considers the common view of judgment, as consisting essentially in an act of comparison through which perceptions are connected in consciousness, with the critical view, according to which an empirical judgment only attains objective validity through the subsumption of the perceptions compared under a pure concept of the understanding. This suggests that the distinction between judgments of perception and judgments of experience is between two *conceptions* of judgment rather than, as initially presented, two kinds of empirical judgment.<sup>41</sup> On this reading, Kant is arguing that, according to the common view (the standard empiricist account), which reduces an empirical judgment to a comparison of perceptions as they occur in a particular consciousness, such a judgment could never lay claim to more than a merely subjective validity, a report of how things seem to a particular observer (or set of observers) under particular conditions. In short, on this picture of judgment, Hume would be correct, and there would be no grounds for distinguishing between mere constant conjunction and a genuine causal connection or for inferring the latter from the former. All one could say is that “however often I and others have perceived this [the stone becoming warm when the sun shines on it]; the perceptions are only usually found so conjoined” (Pro 4: 301n.; 95). It could not be claimed that the sun actually *warms* the stone.

On the alternative conception of judgment, however, which assigns a constitutive function to the pure concepts, this unwelcome result is avoided. These concepts generate objectively valid judgments by serving as rules through

which perceptions are connected in a manner that holds not simply for a particular consciousness but universally (for consciousness in general). The core idea, which is developed in greater detail in the *Critique*, is that to relate one's representations to an object just is to think them as unified in a way that is independent of one's perceptual state or any other subjective factors and, therefore, holds for everyone. In other words, the thought of the objective validity of such a unification is equivalent to the thought of its universality and necessity, which is why Kant claims that "objective validity and necessary universal validity (for everyone) are interchangeable concepts" (Pro 4: 298; 93). But thinking one's representations as unified in this way consists in thinking them as ordered in accordance with a rule that itself expresses such necessary universal validity.<sup>42</sup> Since the pure concepts are precisely these rules, it follows that the objective validity of our judgments is crucially dependent upon them.

By thus interpreting the *Prolegomena's* distinction between judgments of perception and judgments of experience as one between two conceptions of judgment, we avoid any conflict with the *Critique*. The problem, however, is that this reading does not accord completely with the text of the *Prolegomena*, because Kant also takes the distinction to characterize successive stages in a cognitive process in which the understanding proceeds from merely subjectively valid "seemings" or "perceivings" to objectively valid cognitions. Indeed, early in his discussion, Kant explicitly states,

All our judgments are at first mere judgments of perception; they hold only for us, i.e., for our subject, and only afterwards do we give them a new relation, namely to an object, and intend that the judgment should also be valid at all times for us and for everyone else. [Pro 4: 298; 92]

In viewing judgments of perception as the first stage in a cognitive process leading to judgments of experience, Kant treats them as actual judgments (or at least "preliminary judgments" (*vorläufige Urteile*), which does seem to conflict with §19 of the *Critique*.<sup>43</sup> Nevertheless, as we shall soon see, the *Prolegomena's* account of empirical cognition as involving a move from perception to experience has an echo in §26 of the B-Deduction, albeit not in terms of a distinction between two kinds of judgment. Accordingly, it is there (rather than in §19) that we must look in order to gain a proper understanding of the compatibility of the *Prolegomena's* account with that of the B-Deduction.<sup>44</sup>

Since we shall also see that the central claim of §26 is that perception is itself governed by the categories, this seems to bring us back to our original problem. A reconciliation is possible, however, if, as seems reasonable, the *Prolegomena's*

neglect of this crucial point is attributed to its particular goal and method rather than to any philosophical confusion or doctrinal change on Kant's part. First, the focus on judgment and the conditions of its objective validity may easily be understood in the light of the analytic method adopted in that work. Assuming the actuality of judgments possessing this normative property, Kant inquires into the conditions of their possibility, which he explicates by showing the judgmental function of the categories. Second and most important, for Kant to have gone beyond that and attempted to explain the role played by the categories in perception would have required the introduction of the transcendental function of the imagination. But the inclusion of this topic would not only have greatly complicated the account, thereby compromising the suitability of the *Prolegomena* as "preparatory exercises" (Pro 4: 261; 58); it would also have focused attention on precisely those features of his account emphasized by Kant's critics, namely, its inherent obscurity and its apparent subjectivism. Accordingly, it is understandable why Kant should have remained silent on this important matter in the *Prolegomena*, even though doing so had the potential (unfortunately fully realized) of leading to serious misunderstanding.<sup>45</sup>



### B. Subjective Unity in the Critique

As I have already indicated, the B-Deduction's contrast between the objective and subjective unities of consciousness should be seen as a change of subject rather than a correction of the *Prolegomena*'s distinction between judgments of perception and judgments of experience. The change is from a reflection on the conditions under which a *judgment* possesses objective validity to one on the conditions under which a unity of consciousness constitutes a judgment.<sup>46</sup> But just as the problem in interpreting the *Prolegomena* stems largely from the conception of a judgment of perception, so the present problem concerns understanding the conception of a non-judgmental unity of consciousness, that is, a subjective, merely *de facto* unity, to which Kant attributes (as he does to judgments of perception) subjective validity.

At the heart of the problem lies an ambiguity in the conception of subjective unity. It can mean either a unity *in consciousness*, that is, a unity or order possessed by a set of representations, or a unity *for consciousness*, that is, a unity of which a subject is in some sense aware (though not as objective), which therefore corresponds to a certain (purely subjective) form of awareness. Not only does Kant understand such a unity in both ways, but the unity comes in two distinct forms. The first of these is the unity through which the data for an act of judgment are empirically given, which Kant describes as a "determination of



inner sense" (B139). This refers to the order in which perceptions occur in inner sense or, equivalently, the "order of apprehension," for example, the order in which I successively apprehend the parts of a house. Such an order is "subjective" in the sense that I do not *judge* that these successively perceived parts objectively succeed one another in time. The other is the unity produced through association. As examples of this unity Kant cites the distinct connections that different subjects might make with the representation of the same word (B140) and the association of carrying a body and feeling a pressure of weight, which is distinguished from the judgment "It, the body, is heavy" (B142).

Although these forms of subjective unity have quite different causal ancestries, they share a non-representational status, because they are not unities *through which* something is represented or "taken" as an object. To associate *A* and *B* is not of itself to represent them as combined in an object, though it can become the basis (albeit an inadequate one) for such a judgment. Similarly, unities *in consciousness* can become unities *for consciousness* by being represented reflectively as objects of inner sense. But in that case the subject presumably forms a judgment about its state, which, like any judgment, is objectively valid and stands under the categories.

If Kant had understood by a subjective unity of consciousness nothing more than these two forms of non-representational unity, the distinction between the objective and subjective unities of consciousness would have been non-problematic. Thus, it is tempting to read his account of subjective unity as merely affirming this, just as it was tempting to take the *Prolegomena's* distinction between judgments of perception and judgments of experience as holding between two conceptions of judgment.

Once again, however, the text does not permit such a reading. The problem is that it also refers to a unity for consciousness and, therefore, a form of awareness that is not "objective" in the Kantian sense, since it does not constitute an actual judgment. This is the empirical unity of apperception, which Kant introduces in passing as a contrast to the transcendental unity. Only the latter, Kant points out, is objectively valid; whereas "the empirical unity of apperception, which we are not assessing here, and which is also derived only from the former under given conditions *in concreto*, has merely subjective validity" (B140).

Assuming that Kant here means by empirical apperception what he usually does, namely, an immediate awareness of the contents of consciousness (as opposed to a judgment about these contents), then the unity of this apperception is something *for*, as well as *in*, consciousness, even though it lacks any objective

validity.<sup>47</sup> Moreover, although much remains obscure at this point, it is at least clear that our initial understanding of the contrast between the objective and subjective unities was too simplistic. Since empirical apperception is a mode of awareness (empirical consciousness) distinct from a judgment about the contents of inner sense (which, as a judgment, is objectively valid), we must understand the subjective unity of consciousness to include a subjective mode of awareness as well as the non-representational unities in consciousness described above.<sup>48</sup>

A further complication is introduced by Kant's claim that the empirical unity of apperception is derived from the transcendental or objective unity "under given conditions *in concreto*." Assuming that by the former Kant intends a form of awareness, this must mean that, though the content of such consciousness is determined by contingent empirical factors (for example, the order in which perceptions happen to occur in empirical consciousness), its form as a mode of consciousness is subject to the transcendental conditions of unity.<sup>49</sup> Although this is precisely what Kant's analysis of apperception leads one to expect, it fails to explain how such a non-objective form of consciousness is possible. But rather than attempting to do so, Kant casually dismisses the problem by remarking that here he is not assessing the empirical unity of apperception.

In spite of differences in the issues, the present problem is connected in an interesting way with the one encountered in the *Prolegomena*. Just as we saw there that the initial understanding of the distinction between judgments of perception and judgments of experience as two theories of judgment breaks down because Kant regards the former as a distinct form of awareness (if not of judgment), we now see that our initially promising interpretation of the distinction between objective and subjective unity is vitiated by the fact that the latter includes (as one of its modes) that very form of awareness. This also suggests, however, that the problem may be dealt with in a similar way, namely, by focusing on the context in which the distinction is introduced.

In the case of the *Prolegomena*, we saw that it was the analytic method and limited scope of his concern that led Kant to ignore the issue. By contrast, in the B-Deduction it is the two-step mode of proof, based on a synthetic procedure, that leads to a similar result. By requiring him to abstract from the cognitive contribution of sensibility (beyond its being the generic source of data) in the first part of the argument, Kant's methodology precludes any consideration of the empirical elements of cognition and their connection with the categories. Thus, as we shall see, it is only in §26, in connection with the empirical

synthesis of apprehension and its role in perception, that Kant introduces the topic of empirical consciousness. We shall also see that it is the introduction of the transcendental synthesis of the imagination in §24 that first puts us in a position to understand the dependence of empirical on transcendental apperception. Consequently, we must turn to these sections not only to complete our analysis of the argument of the B-Deduction but also to gain a deeper understanding of empirical apperception and its subjective unity.

### III. IMAGINATION, PERCEPTION, AND EXPERIENCE

The second part of the B-Deduction brings human sensibility and its *a priori* forms back into the picture. It thus may be viewed as an attempt to synthesize the results of the Metaphysical and first part of the Transcendental Deductions with those of the Transcendental Aesthetic. Or, more precisely, it may be seen as a first step in such a synthesis, which encompasses the rest of the Analytic.

Kant describes the task in §21 as to show “from the way in which the empirical intuition is given in sensibility that its unity can be none other than the one that the category prescribes to the manifold of an intuition in general” (B144–45). In §26 he characterizes it more expansively as explaining “the possibility of cognizing *a priori* through categories whatever objects may come before our senses, not as far as the form of their intuition but rather as far as the laws of their combination are concerned, thus the possibility of as it were prescribing the law to nature and even making the latter possible” (B159–60). Whether these characterizations are equivalent is a matter that requires further discussion. For the present, we need only note that a second step is necessary on either formulation because the unity that the category prescribes in a judgment is a purely conceptual one. It governs the *thought* or conceptual unification of a given manifold of appearance, but it does not guarantee that the latter will be unifiable in the manner specified in the category, in which event the categories would remain “mere forms of thought without objective reality” (B148).

Kant’s strategy for demonstrating that the categories are more than empty thought forms turns on connecting them with the imagination. The argument consists of two steps, which are separated by a discussion of the relation between inner sense and apperception.<sup>50</sup> First, Kant links the unity of apperception, and with it the categories, to time as the form of inner sense (§24). This linkage is based on the connection of both with the transcendental synthesis of the imagination. He then relates the categories to the content of inner sense,



namely, perception or empirical intuition (§26). The focal point of the latter step is an analysis of the empirical synthesis of apprehension. By arguing that this synthesis is conditioned by the transcendental synthesis of the imagination, Kant contends that it, too, is necessarily subject to the categories. This step is complicated, however, by the fact that, after connecting the categories with perception, Kant proceeds without further argument to affirm their status as conditions of experience, defined as “cognition through connected perceptions” (B161). Accordingly, the question of the connection between perception and experience, which first arose in the *Prolegomena* in terms of the distinction between judgments of perception and judgments of experience, must be reconsidered in the light of this new account of perception as itself categorially determined. But before addressing this complex argument, it is necessary to consider briefly Kant’s conception of the imagination and its cognitive function, as well as some of the ways in which these topics are viewed in the literature.

### A. The Imagination and Its Cognitive Function

In introducing the imagination in the Metaphysical Deduction, Kant characterizes it as that “blind though indispensable function of the soul [or understanding],<sup>51</sup> without which we would have no cognition at all, but of which we are seldom even conscious.” This function is synthesis, whereas that of the understanding is “to bring this synthesis to concepts,” thereby first producing cognition “in the proper sense” (A78 / B103). Setting aside the notorious obscurity of this characterization, its basic point is simply that the imagination is an essential ingredient in cognition, without itself yielding cognition “in the proper sense.” This is a core Kantian thesis and is affirmed in both versions of the Transcendental Deduction.

Although never explicitly defined, in the A-Deduction the imagination is on stage from the opening scene and is accorded the dignity of being a separate cognitive power mediating between sensibility and understanding (A124). Kant there also provides an extensive, if diffuse, account of its cognitive function, which is epitomized in the remark, “No psychologist has yet thought that the imagination is a necessary ingredient of perception itself” (A120n). By contrast, in the B-Deduction the imagination makes a belated entrance in §24 under the guise of the figurative synthesis (*synthesis speciosa*), where it is defined in fairly traditional terms to which no empiricist would object as “the faculty for representing an object even **without its presence** in intuition” (B151). But

rather than viewing it as a distinct faculty, Kant now links it closely to the understanding. Thus, he initially describes the transcendental synthesis of the imagination as an “effect [*Wirkung*] of the understanding on the sensibility” (B152), and claims that it is the understanding “under the designation of a transcendental synthesis of the imagination” that determines inner sense (B153). Indeed, at one point he maintains that “[i]t is one and the same spontaneity that, there under the name of imagination and here under the name of understanding, brings combination into the manifold of intuition” (B162n.).



This connection of the imagination with perception, on the one hand, and with the understanding, on the other, has caught the attention of some prominent philosophers, who see it as an anticipation of familiar Wittgensteinian views. Thus, Strawson, though adopting a literal reading of Kant’s definition of the imagination as an image-making faculty, attributes to it an essential role in perceptual recognition.<sup>52</sup> As he puts it at one point, “[T]he visual experience [of an object as an object of a certain kind, or as a particular object of that kind] is *irradiated* by, or *infused* with, the concept; or it becomes *soaked* with the concept.”<sup>53</sup> Similarly, Wilfred Sellars, though paying considerably greater attention to the texts and insisting on a sharp distinction between imagining and imaging, argues that perception for Kant is already in a sense theory laden.<sup>54</sup> According to Sellars, it is because of the imagination’s role in perception that intuition (as distinct from images and sensations) has a categorial form and even “embodies a proto-theory of a world.”<sup>55</sup>



Although these attempts to rehabilitate Kant’s conception of the imagination are of considerable philosophical interest and make contact with deep Kantian themes, they are somewhat problematic as interpretations of Kant. In the case of Strawson, the problem is twofold. First, in spite of its accord with the letter of Kant’s definition, Strawson’s understanding of the Kantian imagination as primarily a capacity for forming images brings it closer to familiar empiricist views than it actually is.<sup>56</sup> Second, and in apparent contradiction to the first, he over-intellectualizes the imagination by conflating its recognitional role with that which Kant assigns to concepts.

Since Sellars explicitly distinguishes between imaging and imagining, he avoids the first problem and is keenly aware of the second. Thus, he distinguishes between perceptual takings, which are attributed to the imagination and provide a perceiver with the subject terms for judgment, and judgment proper, which involves explicit predication.<sup>57</sup> Elsewhere he draws essentially the same distinction in terms of a contrast between the “minimally conceptual”

character of the imaginative synthesis through which we apprehend a particular, for example, this cube or this white thing, and the fully conceptual representation expressed in a judgment about it.<sup>58</sup>

Nevertheless, Sellars too comes perilously close to over-intellectualizing the Kantian imagination, albeit for complex reasons that cannot be fully considered here.<sup>59</sup> In spite of the close affiliation between the imagination and the understanding affirmed in the B-Deduction, Kant retains the division of labor between them: the task of the imagination is to synthesize and of the understanding to bring this synthesis to concepts. Moreover, this suggests that the activity of the imagination is not itself directly conceptual in even the “minimal” sense indicated by Sellars. Instead, it is more appropriately described as “proto-conceptual.” Although the distinction between being “minimally conceptual” and “proto-conceptual” may seem overly subtle, it is essential for a proper interpretation of Kant’s complex doctrine. The basic point is that the imagination has the task of unifying the sensible data in a way that *makes possible* its subsequent conceptualization, without itself *being* a mode of conceptualization.

In fact, the way in which the imagination performs its task at the level of perception is already indicated by Sellars when, in a phenomenological vein, he points out that the products of the imaginative synthesis, which he terms “image-models,” are perspectival or “point-of-viewish” in nature.<sup>60</sup> In other words, perception (at least visual perception) is always from a point of view, with the object presenting itself as having aspects that are not actually “given.” This presentational function, which is necessary for the subsequent cognition of the object as being of a certain kind, shape, size, and so forth, is the work of the imagination.

The basic idea underlying this account is brought out in a somewhat clearer manner by J. Michael Young.<sup>61</sup> Like Sellars, Young denies that the imagination is to be viewed primarily as a capacity for “imaging.” And like Strawson, Sellars, and those who approach Kant through the hermeneutical tradition, he locates its function in interpreting or construing perceptual data.<sup>62</sup> Moreover, again in agreement with Sellars, Young emphasizes the perspectival nature of such interpretation. Thus, for him imaginative interpreting consists in taking something as “other or more than it is perceived as being,” which is his non-imagistic gloss on Kant’s definition of the imagination.<sup>63</sup>

Young differs from Sellars, however, in insisting on the pre-linguistic, pre-conceptual nature of this perceptual interpreting. Thus, while underscoring the rule-governed nature of the imaginative synthesis, he preserves its non-concep-

tual character by interpreting its relation to the understanding in terms of Kant's distinction between proceeding according to a law and according to the conception of a law (Gr 4: 413; 66). Substituting 'rule' for 'law', Young's basic point is that the imagination proceeds in the former manner and the understanding in the latter. In other words, whereas the imagination remains "blind" in the sense that its rule-governed interpretive activity is not self-consciously performed, which is why it does not amount to cognition "in the proper sense," the understanding, proceeding according to the *conception* of a rule, reflectively takes some content as falling under a concept, thereby forming a judgment that makes a normative demand on the agreement of others.<sup>64</sup> Let us assume, then, the correctness, at least in its broad outlines, of this view of the Kantian imagination as a proto-conceptual, interpretive faculty, distinct from a capacity to produce images, on the one hand, and to form judgments, on the other. And now let us see whether it helps to illuminate Kant's progressive argument in the second part of the B-Deduction linking the categories with appearances by connecting both with the imagination.

## B. The Transcendental Synthesis of the Imagination

As I have already indicated, the first step occurs in §24, where Kant introduces the figurative synthesis or "*synthesis speciosa*" and distinguishes it from the purely intellectual synthesis attributed to the understanding. In this context, Kant makes two claims concerning the former on which further progress of the argument depends: (1) that this synthesis, under the name of the transcendental synthesis of the imagination, has both an *a priori* dimension and a transcendental function in the determination of time as the form of inner sense; and (2) that since this function must accord with the conditions of the synthetic unity of apperception, it is subject to the categories (B151–52). Clearly, if this two-part thesis can be demonstrated, we would be considerably closer to seeing how the categories serve to structure perception, thereby attaining objective reality. Moreover, even though it cannot be claimed that Kant himself actually establishes either part of this thesis, he does supply the necessary materials for constructing the relevant arguments.

To begin with, there is the previously noted definition of the imagination. A capacity to represent what is not present in intuition is clearly necessary for the representation of time as described in the Aesthetic. The essential point is that each determinate extent of time is intuited as a portion of a single all-inclusive time, which is itself represented as an infinite given magnitude. It follows from

this that the awareness of a determinate time involves the awareness of it as a portion of this single time. But, as with space, this whole is not itself actually given in intuition as an object. Time is, as it were, given only one moment at a time. Consequently, in order to represent a determinate time, we must be able to represent past and future times that are not “present” and ultimately the single time of which they are parts. But this requires a contribution of the imagination.

The question of how this contribution is to be understood is more difficult. Even though in the Schematism Kant characterizes time as “[t]he pure image . . . for all objects of our senses in general” (A142 / B182), it is evident that the representation of this single, all-inclusive time cannot literally be a matter of forming an image. But it likewise seems problematic to characterize such a representation as an “interpretation,” since it apparently lacks a feature that we normally require of anything that counts as such, namely, the availability of alternatives. Nevertheless, Kant’s illustration of the thesis that an imaginative synthesis is required for the representation of time does indicate a sense in which it may be regarded as something like an interpretation. Thus, after noting in the case of space that we cannot think a line without drawing it in thought and a circle without describing it, or represent the three dimensions of space without placing three lines perpendicular to each other at a point, Kant remarks that we cannot represent time

without, in **drawing** a straight line (which is to be the external figurative representation of time), attending merely to the action of the synthesis of the manifold through which we successively determine the inner sense, and thereby attending to the succession of this determination in inner sense. [B154]

This suggests not only that the imaginative representation of time presupposes a figurative synthesis (which applies to space as well) but also that it involves an interpretation of the line as an image of time. The need for such an interpretation stems from the fact that inner sense has no manifold of its own. Since this is the case, the materials for the representation of time must be provided by outer sense, which means that time is necessarily represented or “interpreted” in spatial terms. In Young’s terms, the line is taken “as something more or other than it is perceived as being.” Indeed, it may even be said to leave room for an alternative interpretation, since one is always free to abstract from the generative process and consider merely its product, namely, the line itself, purely as a determination of space.



For the purposes of the Deduction, however, the main question is why the imaginative synthesis that determines time must conform to the categorial requirements of the understanding. Although the analysis of the representation of time in terms of the drawing of a line suggests that the category of quantity is involved, it does not seem possible to generalize from this to the remaining categories. Moreover, it is precisely the figurative nature of the imaginative synthesis that renders its conformity to the purely conceptual requirements of the understanding problematic. Accordingly, our original question concerning the necessity of such conformity on the part of the sensibly given reappears in connection with the imagination.

Unfortunately, Kant's treatment of this issue is extremely perfunctory. Instead of providing an argument, he simply asserts that the imaginative synthesis is an expression of the spontaneity of thought, that it determines inner sense *a priori* in respect of its form, and that this determination is in accord with the unity of apperception (B151–52). The latter point is certainly germane. If the determination of time is in accord with the unity of apperception, then, given the argument of the first part of the Deduction, it must also be in accord with the categories. But, again, the problem is to see why this figurative synthesis must accord with the conditions of the synthetic unity of apperception governing discursive thought.

As a first step in dealing with this problem, it should be noted that it follows from Kant's account of apperception that the *representation* of time through the drawing of a line must conform to the conditions of its synthetic unity. Not merely the intuitive representation of a line as such but also the "interpretation" of its successive synthesis as the pure image of time presupposes a single subject conscious of its identity throughout the generative process. In short, even though, as we saw in chapter 5, time itself is not a synthetic unity composed of preexisting parts, its determinate representation requires a synthetic unity of consciousness, which brings this representation under the categories.<sup>65</sup>

Nevertheless, this is only one half of the story. Although all imaginative *representing* must conform to the conditions of the unity of apperception, what is *represented* (time and space) is determined by the forms of human sensibility. In other words, the necessity of representing all appearances in a single time and space is not a demand imposed on sensibility by the understanding but is rather required of the latter by the former.<sup>66</sup> This is also why the second part of the B-Deduction is not a trivial move from the conditions of the representation of a manifold of intuition in general to the conditions of the representation of the

pure manifold of time. The introduction of human sensibility and its *a priori* forms brings with it an element that cannot be accounted for in terms of the purely conceptual requirements of the understanding.

The partial dependence of thought on conditions imposed by the nature of sensibility is illustrated by a passage that is frequently thought to indicate precisely the opposite, namely, the much discussed footnote in §26. Even though this note was briefly discussed in chapter 5 in connection with the analysis of the “givenness” of space in the Aesthetic, it is useful to quote it again in the present context:

Space, represented as **object** (as is really required in geometry), contains more than the mere form of intuition, namely the **comprehension** [*Zusammenfassung*] of the manifold given in accordance with the form of sensibility in an **intuitive** representation, so that the **form of intuition** merely gives the manifold, but the **formal intuition** gives unity of the representation. In the Aesthetic I ascribed this unity merely to sensibility, only in order to note that it precedes all concepts, though to be sure it presupposes a synthesis, which does not belong to the senses but through which all concepts of space or time first become possible. For since through it (as the understanding determines the sensibility) space and time are first **given** as intuitions, the unity of this *a priori* intuition belongs to space and time, and not to the concept of the understanding (§24). [B160–61, note]

If, as is usually done, one focuses on the first sentence, the note seems to suggest a fundamental revision of the theory of sensibility outlined in the Aesthetic. Rather than simply being “given” as the initial picture suggested, the actual intuitions of space and time as “formal intuitions” are now seen as constituted by a synthetic activity. In the most radical version of this reading, space and time become *entia imaginaria*.<sup>67</sup> If one considers the note as a whole, however, a quite different picture emerges. First, Kant strongly suggests that the apparent discrepancy between the present account, which emphasizes the necessity of the imaginative synthesis, and that of the Aesthetic, which is silent on the matter, is to be attributed to his manner of presentation rather than to a change in doctrine. Since Kant’s concern in the Aesthetic was to isolate sensibility in order to determine its particular contribution to cognition, he had to ignore at that point the necessity of a synthesis for determinate representations of space and time (formal intuitions). Accordingly, this methodologically necessary omission is now being rectified. Second, when Kant says that in the Aesthetic he attributed this unity to sensibility “in order to note that it precedes all concepts,” he is not denying the dependence of the imaginative synthesis on the categories. As the text makes clear, the concepts Kant there had in mind

were those of space and time, not the categories.<sup>68</sup> Third, when Kant states at the end of the note that “the unity of this *a priori* intuition belongs to space and time, and not to the concept of the understanding,” he is neither retracting the initial claim that the representation of a unified space or time results from a synthesis, through which they are “first given as intuitions” (formal intuitions), nor contradicting the dependence of this unifying synthesis on the categories. Once again, the point is rather that the imaginative synthesis is itself constrained by space and time as forms of intuition, which is just what §24 requires. Accordingly, it is no accident that Kant refers the reader back to §24.

### C. The Synthesis of Apprehension

Although the connection of the categories with the forms of sensibility through the transcendental synthesis of the imagination is undoubtedly the pivotal step of the argument, it does not of itself suffice to secure the goal of the Deduction. This requires demonstrating that the categories stand in a necessary connection with *empirical* intuition. Kant attempts to achieve this goal in §26 by linking the categories to the synthesis of apprehension, which he defines as “the composition [*Zusammensetzung*] of the manifold in an empirical intuition, through which perception, i.e., empirical consciousness of it (as appearance), becomes possible” (B160).

As usual, the initial problems here are terminological. Specifically, they concern the relation between ‘empirical intuition’, ‘perception’, ‘empirical consciousness’, and ‘appearance’. Basically, Kant is claiming that empirical intuition (meaning the content thereof) only becomes perception (understood as an actual consciousness of this content) when its manifold is apprehended, that is, synthesized or taken up in a single empirical consciousness. The latter is “empirical” because it is a consciousness of what is immediately present to a particular subject in perception rather than of what is determined to be the case in empirical cognition or a judgment of experience. Thus, in sharp contrast to the *Prolegomena*, Kant now claims that even perception, though no longer viewed as a type of judgment, stands under the categories.<sup>69</sup>

Although establishing this claim is essential to the goal of the Deduction, Kant’s account is quite sparse and lacks the richness that characterizes the discussion in the A-Deduction. Rather than arguing that perception involves an imaginative synthesis or “interpretation” in the sense previously indicated, Kant simply assumes this to be the case and devotes his efforts to showing that it, like the transcendental synthesis of the imagination, is subject to the categories. The argument for this thesis consists of six steps, once again compressed

into a single paragraph (B160–61). Since the first five form a unit, they will be treated together in the present section, saving a discussion of the sixth, which takes the argument in a new direction, for the next.

[Step 1.] We have **forms** of outer as well as inner sensible intuition *a priori* in the representations of space and time, and the synthesis of the apprehension of the manifold of appearance must always be in agreement with the latter, since it can only occur in accordance with this form.

As noted above, Kant here assumes the actuality of a synthesis of apprehension. Taking this for granted, he reminds us that this synthesis must conform to space and time. The point of this reminder is to indicate that whatever turns out to be a necessary condition for the determinate representation of space and time will also be a necessary condition for the apprehension or perception of anything intuited in space and time.

[Step 2.] But space and time are represented *a priori* not merely as **forms** of sensible intuition, but also as **intuitions** themselves (which contain a manifold), and thus with the determination of the **unity** of this manifold in them (see the Transcendental Aesthetic).

This is the passage to which Kant attaches the previously discussed footnote in which he distinguishes between a form of intuition and a formal intuition. The systematic significance of the step (including the note) is to underscore the point (already made in §24) that space and time are not only forms of intuition but also themselves intuitions with a manifold (content) of their own. As such, they can only be represented insofar as their manifold is unified. Although Kant does not refer explicitly to it in either the text or the note, it is clear that the transcendental synthesis of the imagination is the vehicle for this unification, and thus for a determinate representation of space or time.

[Step 3.] Thus even **unity of the synthesis** of the manifold, outside or within us, hence also a **combination** with which everything that is to be represented as determined in space or time must agree, is already given *a priori*, along with (not in) these intuitions, as condition of the synthesis of all **apprehension**.

This step is effectively a combination of the first two. Kant here states that the conditions of the representation of the unity of space or time are also conditions of the apprehension of anything as determined in them. Consequently, anything that is represented as determined in space or time must conform to the conditions of the representation of their unity.

The key conception here is that of representing something as determined in space or time. Although Kant does not explain what he understands by this, it seems reasonable to assume that it involves (in the case of space) perceiving something as having a certain figure and position and (in the case of time) perceiving an event as having a certain duration and place in a temporal order (as before *A*, after *B*, simultaneous with *C*, and so on).<sup>70</sup> This makes explicit the point that is already implicit in step 1. But since we learned in §24, and have seen reaffirmed in step 2, that the determination of space and time is achieved through the transcendental synthesis of the imagination, it also serves to connect the empirical synthesis of apprehension to the latter. The claim that this synthetic unity is given “along with (not in) these intuitions” reflects Kant’s basic principle that the representation of unity requires a synthetic activity and cannot be passively received through sensibility.

[Step 4.] But this synthetic unity can be none other than that of the combination of the manifold of a given **intuition in general** in an original consciousness, in agreement with the categories, only applied to our **sensible intuition**.

Although obviously the key step, since it links the synthesis of apprehension with the categories, Kant once again offers no argument. Instead, he simply asserts that the unity required for apprehension is an application to human sensibility of the unity of the manifold of an intuition in general that is required for apperception, which licenses the claim that the former, like the latter, is governed by the categories. Nevertheless, we can see that this result does follow from the assumption that the transcendental synthesis of the imagination is governed by the categories taken in connection with step 3. At least it follows if we construe this step to assert that the synthesis of apprehension is subject to the conditions of the transcendental synthesis of the imagination. Accordingly, virtually the entire weight of the argument falls on the claim that the latter is governed by the categories.

[Step 5.] Consequently all synthesis, through which even perception itself becomes possible, stands under the categories . . .

By claiming that even the synthesis through which perception becomes possible stands under the categories, Kant draws the conclusion to which the previous four steps are intended to lead. Clearly, however, this follows only if it is assumed that perception already involves representing what is perceived as determined in space and time in the sense previously designated. Confining our-

selves for the present to space, the claim is that to perceive or apprehend something is to take it as having, in addition to its sensory qualities, a certain spatial form and magnitude.

Moreover, it must be kept in mind that this perceptual taking or “interpretation” is the work of the imagination and does not involve a discursive judgment. In other words, we are dealing with a merely “empirical consciousness” and a “subjective unity,” which, according to both the *Prolegomena* and §18 of the *Critique*, supposedly lack categorial determination. In initially considering these texts, I suggested that Kant’s apparent exclusion of a role for the categories was to be understood as an omission rather than a denial and that in both cases this is to be explained on methodological grounds. But there was no attempt at that point to explain what role the categories play at this “proto-conceptual” level. This issue can no longer be evaded, however, because it is essential to the success of the Deduction.

Since it is a matter of perception rather than judgment, the categories cannot here serve as predicates.<sup>71</sup> Accordingly, it is not a matter of subsuming an object that one takes to be a house under the category of substance, but rather of a category functioning as a “rule of apprehension,” directing the empirical synthesis of the manifold of the house in accord with the condition required for the determination of the space it occupies.<sup>72</sup> Consider the first of two illustrations of the function of a category that Kant provides in §26:

Thus if, e.g., I make the empirical intuition of a house into perception through apprehension of its manifold, my ground is the **necessary unity** of space and of outer sensible intuition in general, and I as it were draw its shape in agreement with this synthetic unity of the manifold in space. This very same synthetic activity, however, if I abstract from the form of space, has its seat in the understanding, and is the category of the synthesis of the homogeneous in an intuition in general, i.e., the category of **quantity**, with which that synthesis of apprehension, i.e., the perception, must therefore be in thoroughgoing agreement [B162].

Kant’s point here is that the perception of a house, which tokens a three-dimensional object, is conditioned by the determination of the space it is perceived to occupy. It is not, however, a matter of two distinct syntheses: an *a priori* or transcendental synthesis that determines the space and an empirical one that determines the contours and extent of what is perceived in it. It is rather that these are related as the formal and material aspects of one synthesis. The transcendental synthesis of the imagination is the form of the empirical synthesis of apprehension in the sense that the apprehension or perception of a house

is governed by the conditions of the determination of the space it is perceived to occupy. Otherwise expressed, the empirical synthesis is necessarily subject to a “rule of apprehension.”

In this case, the rule is provided by the category of quantity. But the necessity of the category stems not from any *conceptual* requirement concerning the conditions of judgment but from the conditions imposed on the figurative synthesis by the singularity and homogeneity of space. Since the pure concept of quantity is just the concept of the “synthesis of the homogeneous in an intuition in general” (B162), Kant concludes that the apprehension or perception of a house must be governed by that category.

#### D. Completing the Deduction: Perception and Experience

[Step 6.] . . . and since experience is cognition through connected perceptions, the categories are conditions of the possibility of experience, and are thus also valid *a priori* of all objects of experience.

Although part of the same sentence as the preceding step, this constitutes a distinct step in the argument, because Kant here moves from perception to experience. Moreover, this move, which Kant makes without any argument other than an appeal to the definition of experience as “cognition through connected perceptions,” raises at least two questions that are central to the interpretation and evaluation of the B-Deduction.

The first concerns the function of this move within the project of the second part of the Deduction, which we have seen Kant describe in two ways that are not obviously equivalent: (1) showing “from the way in which the empirical intuition is given in sensibility that its unity can be none other than the one that the category prescribes to the manifold of an intuition in general” (B144–45); and (2) explaining “the possibility of cognizing *a priori* **through categories** whatever objects **may come before our senses** . . . as far as the laws of their combination are concerned, thus the possibility of as it were prescribing the law to nature and even making the latter possible” (B159–60). If the project is understood in the first manner, the goal is arguably accomplished by relating the categories to perception in the way indicated in Steps 4 and 5. Moreover, this also seems sufficient to exorcize the specter that has haunted us from the beginning. On this reading, then, the move from perception to experience is peripheral to the main argument. If one understands the project in the second sense, however, it is necessary to show that the categories are conditions of the possi-

bility of experience, which is the conclusion Kant actually draws. But, then, the move from perception to experience stands in serious need of justification.

The second and closely related question concerns the understanding of the claim that the categories are conditions of the possibility of experience. This can be taken in two ways, depending on the weight one places on the components of Kant's definition of experience. On the one hand, if one focuses simply on the fact that perception is a necessary ingredient in experience, it follows that if Kant has shown that the categories are conditions of the possibility of the latter, he has also shown that they are (necessary) conditions of the former. Thus, the conclusion Kant draws is compatible with the weaker reading of the Deduction as primarily an exercise in transcendental exorcism. On the other hand, if one emphasizes the *difference* between perception and experience, which is both an essential feature of Kant's position and built into the definition, then showing that the categories are conditions of the latter requires showing that they make possible cognitive experience as distinct from pre-cognitive (because merely proto-conceptual) perception. But it is far from clear that the argument up to this point is able to deliver this result.

Unfortunately, the issue is further obfuscated rather than resolved by Kant's second illustration, which involves the category of causality and its role in the perception of the freezing of water. Kant notes that the perception of such an event involves the apprehension of the succession of two states of the water: fluidity and solidity. The problem is to explain how such apprehension is possible, and to this end he writes:

But in time, on which I ground the appearance as **inner intuition**, I represent [to myself] necessary synthetic **unity** of the manifold, without which that relation could not be **determinately** given in an intuition (with regard to temporal sequence). [Br62–63]

Although the sentence is hardly a model of clarity, the basic point is that the apprehension of a determinate sequence in time (the successive states of the water) presupposes the representation of the synthetic unity of time, and thus a synthesis governed by a category. In other words, only as the result of such a determination of time, which is presumably accomplished by the transcendental synthesis of the imagination, is it possible to apprehend a determinate sequence of perceptions in time. The further claim, which Kant does not even attempt to justify, is that the ground of this determination or synthetic unity is the category of causality, "through which, if I apply it to my sensibility, I determine



everything that happens in time in general as far as its relation is concerned.” And, given this, he concludes:

Thus the apprehension in such an occurrence, hence the occurrence itself, as far as possible perception is concerned, stands under the concept of the relation of effects and causes, and so in all other cases. [B163]

Perhaps the most striking feature of this example is the lack of any reference to experience. Rather than claiming that the category of causality is a condition of the *experience* of an event or occurrence, which is what one might expect at this point, Kant continues to speak of apprehension or perception. This suggests that the point of the example is to indicate a parallel between the functions of the categories of quantity and causality. Both are directly involved with perception, the former with that of outer sense and its spatial form, the latter with inner sense and its temporal form (succession). Thus, both categories function as rules of apprehension.

Not only, however, does this conflict with the Second Analogy, which endeavors to show that the category functions as a condition of experience rather than mere apprehension or perception; it also leads to a deeply counter-intuitive reading of the B-Deduction. This is because it puts virtually the entire weight of the argument in its least likely place, namely, that very perceptual consciousness, which, according to both the *Prolegomena* and §18 and §19 of the *Critique*, is a form of awareness that supposedly lacks categorial determination! It seems as if Kant has completely reversed course, now claiming that it is perception that is the prime locus of categorial determination and that the subsequent application of the categories to experience and the objects thereof is nothing more than a consequence of this.

If we consider the problem in the light of the distinction Kant introduces in the second edition between the mathematical and the dynamical categories (B110), which corresponds to the distinction he had already drawn in the first edition between the mathematical and dynamical principles (A161–62), one might say that the argument up to this point has, at most, established the validity of the former but not of the latter. This is because the different category types, like their corresponding principles, have quite distinct epistemic functions. Those of quantity and quality are directly concerned with the intuition of objects and, therefore, with their perception, whereas the latter or, more specifically, the relational categories, are concerned with the existence of these objects in relation to each other in time.<sup>73</sup> But, as we shall see in chapter 9, this

existence is a matter of experience ("cognition through connected perceptions") rather than mere perception, since it involves the thought of a determinate objective temporal order that cannot be identified with the order of the apprehension of perceptions in empirical consciousness. That is why an argument geared to linking the categories to perception (as that of §26 seems to be) cannot account for the epistemic function of *all* the categories.

Whether one understands the problem in terms of the twofold characterization of the goal of the argument or in terms of the distinction between categorial functions, it poses something of a dilemma for the interpreter, since it appears to leave us with two unattractive options. One must either acknowledge that the Deduction is at best only partly successful, establishing merely a role for the categories as rules of apprehension, or conclude that Kant himself was unclear about both the nature and scope of his argument.

Moreover, two initially attractive ways of avoiding this dilemma must be rejected as inadequate. One is to argue that the role of the categories as conditions of experience is already established in the first part of the Deduction; so that the concern of the second part is precisely with their role in empirical consciousness or mere perception, that is, a subjective mode of awareness that does not amount to cognition.<sup>74</sup> If the analysis offered here is correct, however, such a reading is based on a serious misinterpretation of the first part of the Deduction. Even though Kant's concern there was with judgment and the conditions of objective validity, and experience differs from mere perception precisely in involving judgment, Kant cannot be taken to have already shown there that the categories are conditions of experience, or even to have thought that he had done so. That would require connecting the categories with the conditions of human sensibility, which first occurs in the second part.

The other, indeed the usual, approach is to dismiss the problem by insisting that the Deduction is concerned with a general proof of the validity of all of the categories and that a consideration of the specific functions of the distinct categories is reserved for the *Analytic of Principles*. Although correct so far as it goes, such a response is inadequate because the argument as presented does not appear fully to accomplish this task. Presumably, the reason this point is usually neglected is that the issue of differential categorial functions does not arise in connection with the first part of the Deduction, where the concern is with the common intellectual function of the categories as grounds of the objective validity of judgment. Thus, it seems reasonable to assume that there the argument covers all of the categories identified in the *Metaphysical Deduction*.

This does not apply to the second part of the Deduction, however, where the focus is on the *experiential* function of the categories. For if, as turns out to be the case, the different category types have significantly different experiential functions, then the argument must somehow incorporate this point. At the very least, it should not suggest that the desired result may be attained by an argument that is applicable only to one such type.

Unfortunately, this is just what Kant does suggest in §26, which means that the interpretive dilemma cannot be avoided. Its seriousness can be somewhat mitigated, however, by attributing the problem to methodological exigencies rather than philosophical confusion. As has been noted on several occasions, Kant's method sometimes makes it difficult (if not impossible) for him to say everything that needs to be said on a given topic at a particular point in the analysis. This was seen to be the source of the apparent tension between the account of space and time in the Aesthetic as "given" through sensibility and the claim in §24 and §26 that they presuppose an imaginative synthesis. In the case of the *Prolegomena*, it was suggested that it was Kant's analytic method and the quite understandable (in the context) emphasis on the *a priori* conditions of empirical cognition that led him to remain silent on the important role of the categories in perception itself. Similarly, it has been argued that it was Kant's concentration on judgment, as discursive thinking considered in abstraction from its connection with the specific nature of human sensibility, that led to essentially the same result in the discussion of subjective unity in §18 and §19 of the B-Deduction. By the same token, then, it seems reasonable to conclude that it is Kant's focus on perception (a focus that is required by the underlying problematic of the Deduction) that prevented him from doing much more in §26 than gesturing (in a misleading manner) toward the complex issues concerning the role of the categories in the connection of perceptions that constitutes experience proper.

Admittedly, this does not amount to a complete vindication of the B-Deduction. The resulting ambiguities constitute a significant defect in its proof strategy, albeit one that has its counterpart in the A-Deduction.<sup>75</sup> Nevertheless, such an approach indicates that these difficulties do not stem from any intrinsic incoherence in Kant's overall position; rather, they suggest that we must look to the Schematism and the Analytic of Principles (particularly the Analogies) in order to complete the account of the connection between the intellectual and sensible conditions of human cognition that is only begun in the second part of the B-Deduction.<sup>76</sup>

## Chapter 8 The Schematism of the Understanding and the Power of Judgment

Like the Deduction, the Schematism has a long-standing reputation for difficulty and obscurity. And, like the Deduction, this reputation is not undeserved. Unlike the Deduction, however, the Schematism has also sometimes been viewed as superfluous. Thus Prichard, who raises this charge against the whole of the second book of the *Transcendental Analytic*, writes:

We naturally feel a preliminary difficulty with respect to the existence of this second part of the *Analytic* at all. It seems clear that if the first part is successful, the second must be unnecessary. For if Kant is in a position to lay down that the categories must apply to objects, no special conditions of their application need be subsequently determined. If, for instance, it can be laid down that the category of quantity must apply to objects, it is implied either that there are no special conditions of its application, or that they have already been discovered and shown to exist. Again, to assert the applicability of the categories is really to assert the existence of principles, and in fact of just those principles which it is the aim of the *System of Principles* to prove. Thus to assert the applicability of the categories of quantity and of cause and effect is to assert respectively the principles that all objects

of perception are extensive quantities, and that all changes take place according to the law of cause and effect.<sup>1</sup>

A similar objection, albeit with specific reference to the Schematism, has been raised by J. G. Warnock, who maintains that the presence of this chapter in the *Critique* is due entirely to Kant's illicit separation of the possession of a concept from the ability to use it. According to Warnock, the goal of the Deduction is to prove that we possess a certain set of concepts, and if this goal were achieved, there would be no remaining questions about the applicability of these concepts and, therefore, no problem requiring a theory of schematism for its solution.<sup>2</sup> Moreover, Warnock has been at least partially followed by Jonathan Bennett, who, while noting that he and Warnock "disagree about the scope and disreputability of Kant's 'problem'," adds that "We agree, however, that Kant does not solve whatever problem he has."<sup>3</sup> Indeed, he goes so far as to claim, "The incoherence of Kant's problem about category-application is matched by the vacuity of its proposed solution."<sup>4</sup>

These dismissive views of Kant's problem are misguided. The question the Schematism addresses is not *whether* the categories apply to appearances (that question is dealt with in the Deduction) but under what (sensible) *conditions* they can do so. Moreover, that question is both perfectly coherent and largely unaddressed by the Deduction. To know that the concept of quantity is applicable to appearances is not yet to know that the objects to which it applies are extensive magnitudes and that all appearances are such magnitudes. The latter conclusion requires the additional knowledge of how this concept is expressed in sensible terms, that is, how it is schematized. The same holds, *mutatis mutandis*, for the other categories.

Similar considerations apply to the objection in its Warnockian form. The purpose of the Transcendental Deduction is not, as Warnock suggests, to prove that we possess a certain set of concepts. On the contrary, it assumes on the basis of the Metaphysical Deduction that we do possess them and that they have at least a "logical use." The question with which it deals is whether they also have an extralogical or "real use," that is, an application to appearances. Once again, the argument, if sound, proves that they do have such a use, but it does not show how and under what specific conditions the particular concepts are to be employed. This is the task of the Schematism. As Kant puts the matter in a *Reflexion*, "The Schematism shows the conditions under which an appearance is determined in respect to a logical function and, therefore, stands under a category" (R5133 18; 392).

In order to understand this difficult chapter, however, it is crucial to recognize that the problem with which it deals is the specification to the particular case of the categories of a general problem regarding the conditions of concept application. Unfortunately, Kant obscures this by the way in which he presents the problem in the Schematism, which suggests that nothing analogous to a schematism is required for concepts other than the categories. But he does point to the broader problem both in the brief Introduction to the second book of the *Analytic*, where he discusses the function of the power of judgment (*Urteils-kraft*), considered as a distinct cognitive faculty, and in two paragraphs within the Schematism itself, where he considers the need for schemata in general terms that apply also to mathematical (pure sensible) and empirical concepts.

The schematization of the categories poses special problems, however, because of the separation of the sensible and intellectual conditions of human cognition. Indeed, this separation seems to preclude the possibility of synthetic *a priori* judgments that apply the categories to appearances, which explains Kant's focus on the specific problem. Nevertheless, the latter remains a specification (due to the unique nature of the categories) of a general problem concerning the application of all concepts. Consequently, it is best approached by first considering the problem in its generality and then turning to the additional complications introduced by the unique status of the categories.

Operating on this assumption, the present chapter is divided into five parts. The first two deal, respectively, with the general problem of schematization and the special problem of a transcendental schematism. The third examines the chief ways in which Kant characterizes transcendental schemata, both in the Schematism and elsewhere. It contends that the basic way is as transcendental time determinations and provides an analysis of this conception. Building on this analysis, the fourth part considers the issue of the nature and justification of the claim that a particular schema belongs to a given category. It suggests that such claims constitute a special class of synthetic *a priori* judgments (here termed "schema judgments") and explores how such judgments might be justified. Finally, the fifth part discusses the connection between the schemata and the principles. It argues that this connection provides the key to the understanding of the synthetic *a priori* status of the latter.

## I. SCHEMATISM AND JUDGMENT IN GENERAL

A notable peculiarity of the second book of the *Transcendental Analytic*, of which the Schematism constitutes the first chapter, is that Kant provides the

book with two distinct titles: “The Transcendental Doctrine of the Power of Judgment” and the “Analytic of Principles.” Since Kant’s main concern is with “Principles of Pure Understanding,” that is, the set of *a priori* propositions resulting from the predication of the categories of objects of possible experience, it is tempting to attribute this duality to Kant’s attachment to his beloved architectonic. On this reading, then, the presence of the former title is due entirely to Kant’s desire to find a place within transcendental logic for the second member of the trinity of cognitive faculties (understanding, judgment, and reason) distinguished in general logic.<sup>5</sup> This temptation should be resisted, however, since there is a real problem concerning judgment that the Schematism addresses. Moreover, once this problem is recognized, it can be seen that a Transcendental Doctrine of the Power of Judgment is an integral element of the project of which the Analytic of Principles (strictly speaking) is a part.

To begin with, it must be emphasized that Kant’s present concern is with the *power* (or faculty) of judgment (*Urteilkraft*), that is, the ability to judge, rather than with the nature and formal conditions of judgment (*Urteil*). We have seen in the two preceding chapters that the latter is the concern of the Metaphysical and first part of the Transcendental Deductions, where we also saw that judgment is intimately related to understanding because “the business of the understanding is to judge.” But this does not mean that the *ability* to judge is equivalent to the ability to understand.

For Kant, the two are related as the ability to formulate and grasp rules and the ability to apply them to particular cases. Thus, understanding is defined by Kant as the “faculty of rules” and judgment as the “faculty of **subsuming** under rules” (A132 / B171). Kant’s concern is therefore with the conditions of the successful exercise of the latter capacity. Clearly, these conditions cannot be identified with the categories. The latter, as we have seen, are epistemic conditions in the sense that they provide the rules validating a synthesis of representations. As such, they ground the normative claim of such a synthesis (or judgment) on the agreement of others. Accordingly, they may also be described as “validating conditions.” But just as objective validity is not equivalent to truth, so conformity to these normative demands does not guarantee the truth of a judgment. Thus, there seems to be a need for a distinct condition (or set of conditions) pertaining to judgment.

Here, as elsewhere, the game of chess provides an apt illustration. To understand the game clearly requires grasping its rules and goal. Since these rules determine which moves are legal or “chessly possible” and the goal defines the purpose for which they are made, it is clear that without this knowledge one

would be unable to play chess at all. As any beginner will testify, however, having such rudimentary knowledge is not sufficient to play the game successfully. For the fact that a move is legal does not make it a good move, that is, one that is called for by the particular circumstances. But the knowledge of the latter is, in Kant's terms, a matter of judgment, which involves "subsuming the particular under the universal" (the set of chessly possible moves dictated by the rules of the game).

Given a knowledge of the purpose of the game, one can easily specify a general criterion for the goodness of a particular move, namely, it must improve one's overall position and, therefore, one's chances of winning or at least forcing a draw. Unfortunately, such knowledge, while necessary, is not sufficient. One also needs to know how to apply this principle to particular cases: Is moving my knight to king's rook 5 really going to improve my position? Moreover, the situation cannot be remedied simply by providing further, more specific rules: for example, by avoiding leaving one's queen unprotected or doubling up one's pawns. For one thing, there are always exceptions to such rules, and it is important to recognize them when they arise. For another, these rules do not exhaust the alternatives. Consequently, they do not relieve one of the necessity of determining for oneself what the particular situation requires, that is, of exercising one's judgment. As Kant succinctly puts it, "[A]lthough the understanding is certainly capable of being instructed and equipped through rules, the power of judgment is a special talent that cannot be taught but only practiced" (A133 / B172).

This oft-cited passage and the underlying claim that (on pain of an infinite regress) there can be no rules for applying rules has met with the approval of many philosophers because it is seen as an anticipation of Wittgenstein's influential account of rule following.<sup>6</sup> But if one is to understand the significance of Kant's denial of rules for judgment, it is essential to keep in mind its highly qualified nature. First, its scope is limited to general logic, which is deemed incapable of providing "precepts" (*Vorschriften*) for the power of judgment. In the case of transcendental logic, however, Kant remarks, "[I]t even seems that the latter has as its proper business to correct and secure the power of judgment in the use of the pure understanding through determinate rules" (A135 / B174). And, anticipating the specific contribution of the Schematism, Kant adds that transcendental philosophy has the advantage (shared only with mathematics) "that in addition to the rule (or rather general condition for rules), which is given in the pure concept of the understanding, it can at the same time indicate *a priori* the case to which the rule ought to be applied" (A135 / B174–75). In



other words, transcendental philosophy (through its designation of the schemata of the categories) and mathematics (through the construction in pure intuition of its “pure sensible concepts”) *can* specify *a priori* the conditions of their rules.

Second, with regard to the power of judgment in general Kant denies only that general logic can provide it with precepts (rules of thumb), and from this he concludes that good judgment cannot be taught. But it does not follow from this that judgment either needs no guidance or is incapable of receiving it. On the contrary, Kant insists that it needs to be practiced and that this usually requires the use of examples. Thus, Kant famously remarks that “examples are the leading-strings [*Gängelwagen*] of the power of judgment, which he who lacks the natural talent for judgment can never do without” (A134–35 / B174).

Although the first of these qualifications is central to the understanding of the systematic function of the Schematism, it is convenient to begin with an examination of the second. Admittedly, Kant denies that examples are strictly necessary, since someone fortunate enough to possess a sufficient natural talent for judgment can do without them. But this points to the question of what examples actually do for judgment and why they (unlike schemata) might be dispensable.

Let us return to the analogy with chess, where the examples might be the classical games of past masters. Clearly, studying such examples is an excellent way to sharpen one’s power of chess judgment. But, equally clearly, it cannot be a matter of blind imitation or slavish rule following. Even if one had complete knowledge of these games, it would be necessary to know how to apply the lessons they contain to the present situation. For example, I might recall that in a similar situation Bobby Fischer once turned apparent defeat into victory by the bold sacrifice of a bishop. But my present situation is only similar, and even if the location of the pieces on the board were perchance identical, the opponent would be different.<sup>7</sup> Thus, what is required is an interpretation of the situation as one falling under the strategic principle embedded in the example or, more simply, recognizing what it is an example of. This is where judgment comes in.

The analogy with chess shows both why examples may be highly useful as pointers without being sufficient and why they are not necessary. Presumably, Fischer himself did not adopt his winning strategy by following an example or, if he did, then the same cannot be said of the original chess “genius” who initially invented this strategy. Moreover, though Fischer, unlike the ordinary player, may not have needed an example to discover the winning strategy, he

did need an ability to grasp what the situation required. But, for the reasons previously given, this grasping cannot be understood in terms of the model of appealing to a set of general rules. It is rather a matter of immediately recognizing the universal (the winning strategy) in the particular, which, in Kant's terms, means possessing the schema.<sup>8</sup>

The relevance of this to the Schematism can be clarified by considering the similarity between examples and images. Images share with examples the property of fixedness or determinacy, that is, they are inherently particular. That is why Kant insists in the Schematism that mathematical and empirical concepts must be grounded in schemata rather than images. The case is clear for mathematical concepts. As Kant notes, no image of a triangle could count as an adequate exemplification of the concept, since "it would not attain the generality of the concept," which makes it valid for all triangles (A141 / B180). This can only be accomplished by its schema, which "can never exist anywhere except in thought, and signifies a rule of the synthesis of the imagination with regard to pure shapes in space" (A141 / B180). In other words, in order to be able to judge that a particular figure is a triangle, one needs not simply the concept (that of a figure inclosed by three straight lines) but also its schema.

Although a philosopher of Warnockian persuasion might question this distinction between concept and schema and deny that we could have the former without the latter, this objection can be met by noting the difference between possessing a geometrical concept in the sense of being able to list the properties of the figure and having the ability to construct it. Admittedly, however, the claim seems more problematical in the case of empirical concepts, such as that of a dog (Kant's example). In fact, here Kant is frequently thought to collapse any meaningful distinction between concept and schema.<sup>9</sup> This is because he defines the *concept* of a dog in a manner similar to his characterization of the *schema* of a triangle, namely, as "a rule in accordance with which my imagination can specify the shape of a four-footed animal in general, without being restricted to any single particular shape that experience offers me or any possible image that I can exhibit *in concreto*" (A141 / B180). If, at least in the case of empirical concepts, *both* the concepts and their schemata are to be construed as rules of the imagination, how can there be a principled distinction between them?

The solution to this puzzle is simply that Kant misstates his own position, referring to the *concept* of dog, when he clearly means the *schema*. This is borne out by both the context of Kant's characterization and its obvious similarity to the schema of a triangle. Consider Kant's preceding remark that (in compari-

son with mathematical concepts) “[e]ven less does an object of experience or an image of it ever reach the empirical concept, rather the latter is always related immediately to the schema of the imagination, as a rule for the determination of our intuition in accordance with a certain general concept” (A141 / B180). What is striking here is the suggestion that the gap between concept and image or object is even wider in the case of empirical than in that of mathematical concepts. Presumably, this is because mathematical concepts have only a limited number of possible forms (determinable *a priori*), whereas in the case of empirical concepts the various forms of possible instantiation (for example, types of dog) are indeterminate. And this indicates that schemata are needed for ordinary empirical concepts as well as *a priori* ones.

As Lauchlin Chipman points out in response to Bennett’s denial of any such need, part of the problem lies with Kant’s choice of the concept of dog as his example. Dogs are so familiar that we find it hard to imagine how someone might have the concept and yet fail to recognize a dog when encountering it. But, Chipman suggests, the situation seems quite different if one considers concepts of less familiar kinds of things, such as tadpoles or bone marrow. In such cases, it seems perfectly plausible to assume that someone might have the concepts in the sense of being able to talk intelligently about the production, size, and behavior of tadpoles and the structure and medicinal powers of bone marrow without being able to recognize a specimen of either.<sup>10</sup>

Understanding the possession of a schema as a recognitional capacity also provides the key to understanding the connection between schemata and the imagination, since the capacity in question is essentially *interpretive*. Specifically, it is a capacity to interpret the sensible data as sufficiently instantiating the criteria thought in the concept to warrant the subsumption of the intuition under the concept. As we saw in the last chapter, such sensible interpretation is precisely the function of the imagination. Thus, even in the case of a familiar object, such as a dog, in order to perceive the animal before one as a member of that species, one must be able to process the sensible data in an appropriate way. Indeed, we have seen that this manner of processing is implicit in the perceptual act itself, wherein it takes the form of an implicit “rule of apprehension.” Such an implicit rule is precisely what is meant by the schema of an empirical concept.

This brings us to the complex and highly controversial matter of whether schemata in general (leaving aside for the present those of the categories) should be regarded as rules. On the one hand, there is the fact that Kant explicitly presents them as such: for example, in the characterizations of the schemata

of the mathematical concept of a triangle and the empirical concept of a dog. On the other hand, we have also seen that Kant defines the understanding (not the imagination) as the “faculty of rules” and suggests that the very need for schemata is a consequence of the fact that the exercise of the power of judgment cannot be governed by rules. Thus, it seems that Kant both requires us to, and prohibits us from, regarding schemata as rules.

Nevertheless, this apparent contradiction disappears if we distinguish between two kinds of rule. Since it is equivalent to a concept, one may be termed a “discursive rule,” and the other, which is equivalent to a schema, a “perceptual rule.”<sup>11</sup> As we have seen, to have a concept (even an empirical one) for Kant involves more than having in mind a set of marks that are associated together on the basis of past experience. It requires the taking of these marks as constituting a “analytic unity” and, therefore, as in a sense being necessarily connected. As Kant himself puts it in the A-Deduction, “[T]he concept of body makes necessary the representation of extension, and with it that of impenetrability, of shape, etc.” (A106).<sup>12</sup> This is what enables the concept to function as a predicate in a possible judgment.

Since it guides perception rather than thought, a perceptual rule (or schema) can neither be spelled out discursively in terms of a set of necessary or sufficient conditions nor function predicatively in a judgment.<sup>13</sup> Instead, it functions to process the sensible data in a determinate way, thereby giving one a sense of what to look for or expect on the basis of certain perceptual “clues.” For example, on seeing the front of a house, one naturally expects that it will have sides and a back with appropriate “house-ish” features. Rules of this sort are intimately connected with the perspectival nature of perception and, therefore, with the imagination. Thus, the suggestion is that the schemata of empirical concepts are to be construed as such rules. Moreover, a rule in this sense does not violate Kant’s strictures against rules for the application of rules, since, as non-discursive, it cannot be formulated in terms of a set of necessary and sufficient conditions.

## II. TRANSCENDENTAL SCHEMATA AND TRANSCENDENTAL SUBSUMPTION

Although it was suggested that Kant’s doctrine of a transcendental schematism should be viewed as an attempted solution for a special case of a general problem concerning concept application, his characterization of the problem in the opening paragraphs of the Schematism focuses exclusively on the problem of

subsuming appearances under the categories, that is, of a “transcendental subsumption.”<sup>14</sup> Kant begins by noting that all cases of the subsumption of an object (or concept thereof) under a concept require homogeneity between the concept and what is subsumed under it. To illustrate such homogeneity, Kant appeals to the relation between the geometrical concept of a circle and the empirical concept of a plate. Here the homogeneity is said to be between the roundness that is thought in the empirical concept and that is intuited in the geometrical concept.<sup>15</sup> Then, by way of contrast to this relatively unproblematic situation, Kant famously remarks:

Now pure concepts of the understanding, however, in comparison with empirical (indeed in general sensible) intuitions, are entirely unhomogenous [*ganz ungleichartig*], and can never be encountered in any intuition. Now how is the **subsumption** of the latter under the former, thus the **application** of the category to appearances possible, since no one would say that the category, e.g., causality, could also be intuited through the sense and is contained in the appearance? This question, so natural and important, is really the cause which makes a transcendental doctrine of the power of judgment necessary, in order, namely, to show the possibility of applying **pure concepts of the understanding** to appearances in general. [A136–37 / B176–77]

In spite of Kant’s emphatic statement, many commentators have found this question to be neither natural nor important. In addition to the previously discussed charge of superfluity, critics have challenged its naturalness by focusing on Kant’s appeal to the notion of subsumption. Assuming that Kant takes the term in the way it is used in the traditional theory of judgment, namely, to designate the relation between a class concept and the particulars falling under it, these critics contend not only that it is totally inadequate to capture the relation of representations in a judgment but also that it misconstrues the relation which, according to Kant’s own doctrine, holds between the categories and the sensibly given. As Kemp Smith puts it, this relation is properly one of form and matter, structure and content, not universal and particular.<sup>16</sup>

Although the opening paragraphs of the Schematism are hardly a model of philosophical lucidity, it would be surprising if Kant were as confused as his commentators suggest. First, one can question whether (as his critics assume) Kant intended the example of the relation between the geometrical concept of a circle and the empirical concept of a plate to be taken as an instance of the subsumption of a particular under a class concept. That would require that the concept of a plate (or rather the plate itself) be taken as a member of the class of circles, which is hardly what Kant meant. Rather, as H. J. Paton points out, the key to Kant’s intent lies in the initially puzzling reference to the roundness that

is “intuited in the latter” (the pure geometrical concept of a circle).<sup>17</sup> This indicates that it is the possibility of exhibiting the geometrical concept in pure intuition (constructing a circle) that supposedly explains the homogeneity between it and objects, such as plates, given in empirical intuition, as well as the empirical concepts that are formed by abstraction from the content of such intuition (“thought in the latter”). The homogeneity, in short, is between pure and empirical intuition, not between a class concept and a member of that class. Moreover, only such a reading enables us to make any sense of the contrast that Kant wishes to draw between the geometrical concept and the pure concepts of the understanding, which “can never be encountered in any intuition.”

Nevertheless, this still leaves us with the task of explaining Kant’s troublesome use of ‘subsumption’. To this end, it should be noted that this is Kant’s second use of the term in connection with the problem of the schematism and that each use appears to take it in a different sense.<sup>18</sup> The first is the previously cited passage in which Kant defines the power of judgment as the “faculty of subsuming under rules.” Here Kant is contrasting the latter with the understanding (the faculty of rules), and the term clearly has its judgmental sense of determining whether particulars fall under universals (general rules). This does not, however, mean that Kant is abandoning his underlying conception of judgment. His intent is not to suggest that the *act of judging* can be adequately analyzed in terms of subsumption; it is rather to call attention to a set of synthetic *a priori judgments* (the Principles of Pure Understanding), which, unlike ordinary judgments of experience, do not merely make use of the categories but actually subsume all appearances under them. Kant’s concern is with the conditions of the possibility of such judgments.

Within the Schematism itself, however, Kant makes it clear that he is using ‘subsumption’ as a synonym for ‘application’.<sup>19</sup> Moreover, in so doing, he now takes it in its syllogistic sense. This is intended to clarify the special problem involved with the application of the categories to appearances.<sup>20</sup> As the text indicates, the problem in this form does not arise in the case of other concepts, since (as the circle-plate example illustrates) they may readily be understood to stand in a relation to intuition. In other words, even though concept and that to which it is applied are still heterogeneous as universal and particular (the traditional Platonic problem), such concepts “are not so different and heterogeneous” from the particulars to which they are applied as to warrant a special treatment. But such a treatment is necessary in the case of the categories because they are not merely different and heterogeneous from appearances; they

are “entirely unhomogenous [*ganz ungleichartig*], and can never be encountered in any intuition.”

Kant sketches his views on syllogistic subsumption in both the *Critique* and the various versions of his *Lectures on Logic*. As he puts it in the *Critique*, inferring or “judging mediately” takes place “through the subsumption of a condition of a possible judgment under the condition of something given.” The “given judgment” is the universal rule that functions as the major premise (“Everything composite is **alterable**”). The minor premise is characterized as “the subsumption of the condition of another possible judgment under the condition of the rule” (“Bodies are composite”). The conclusion is the “mediate judgment,” which results from the application of the rule to the subsumed case (“Consequently, bodies are alterable”) (A330–31 / B386–87). The crucial term here is ‘condition’, since in the minor premise it is the “condition of another possible judgment” (bodies) that is subsumed under “the condition of the rule” (not the rule itself). Kant also tells us that this rule “says something universal under a certain condition.” In Kant’s example this condition is the quality of being composite. The minor premise then asserts that this condition is met in the case of bodies, which licenses the conclusion that all bodies are alterable. The condition of the rule is thus the middle term, the “third thing” that connects the universal rule with the particulars to which it is applied in the conclusion.

Although Kant hardly wished to construe the application of the categories to appearances as a bit of syllogistic reasoning, the analogy underscores the particular problem of understanding how such application is possible. As already noted, the problem is the total heterogeneity of the two elements to be brought into connection, which is due to the fact that, as derived from the very nature of the understanding, the categories are unique among concepts in having no *direct* relation to intuition. Yet, as the Transcendental Deduction demonstrates, they *do* relate to intuition and, therefore, to appearances. Thus, the purpose of the analogy is to suggest that, just as in the case of syllogistic reasoning, the connection between the rule expressed in the major premise and the instance to which it is applied in the conclusion is established only by means of the subsumption of the instance under the condition of the rule. In the present case, where the categories are the universal rules, there is need for some analogue of the “condition of the rule,” or the middle term of the syllogism, under which appearances can be “subsumed.” This analogue will turn out to be the transcendental schema, that infamous “third thing,” which makes possible the mediation of category and appearance.

### III. TRANSCENDENTAL SCHEMATA AS TRANSCENDENTAL TIME-DETERMINATIONS

These considerations raise the question of what, if anything, is capable of playing the role required of a transcendental schema. Unfortunately, rather than providing a direct, unambiguous answer Kant gives us several. Limiting ourselves to the Schematism and ignoring minor variations, we find the following characterizations of a transcendental schema:

1. As a "third thing" or "mediating representation," which "must stand in homogeneity with the category on the one hand and the appearance on the other and which makes possible the application of the former to the latter." Kant also claims that such a representation must be "pure" (nonempirical), and though "**intellectual** on the one hand . . . **sensible** on the other" (A138 / B177).
2. As a "transcendental time-determination," which is "homogeneous with the **category** (which constitutes its unity) insofar as it is universal and rests on an *a priori* rule" and "with the **appearance** insofar as time is contained in every empirical representation of the manifold" (A138–39 / B177–78).
3. As a "formal and pure condition of the sensibility, to which the use of the concept of the understanding is restricted" (A140 / B179).
4. As "only the pure synthesis, in accord with a rule of unity according to concepts in general, which the category expresses" (A142 / B181).
5. As "a transcendental product of imagination, which concerns the determination of the inner sense in general, in accordance with conditions of its form (time) in regard to all representations, insofar as these are to be connected together *a priori* in one concept in accord with the unity of apperception" (A142 / B181).
6. As "nothing but *a priori* **time-determinations** in accordance with rules" (A145 / B184).
7. As "the true and sole conditions for providing them [the pure concepts of the understanding] with a relation to objects, thus with **significance**" (A146 / B185).
8. As "really only the phenomenon, or the sensible concept of an object, in agreement with the category" (A146 / B186).

In spite of significant differences between them, a brief canvass of these characterizations indicates that the primary one is (2) "transcendental time-deter-



mination,” which Kant introduces as the solution to the problem of the possibility of a transcendental subsumption. Although Kant does not provide anything like a definition of a transcendental time-determination, in depicting its mediating function he does tell us a fair amount about what he takes one to be. Specifically, he tells us that “the category constitutes its unity” and that it “rests on an *a priori* rule.” In addition, we already know that as a determination of time it must be connected with sensibility and, therefore, with the imagination and its figurative synthesis. Putting these together, we may preliminarily characterize a transcendental time-determination as a rule-governed product of the figurative synthesis (a “transcendental product of imagination”), which exhibits in a non-discursive manner the form of unity conceptually expressed in a category. Not only does such a characterization meet the conditions required of a transcendental schema, with the possible exception of (8), it is at least compatible with the others. Thus, (1) and (7) are concerned with what a transcendental schema does rather than with what it is, while (6) seems a near synonym. This leaves us with (3), (4), and (5), which emphasize, respectively, its function as an *a priori* condition of sensibility, its connection with inner sense, and its source in the imagination, all of which apply to a transcendental time-determination.

Moreover, it cannot be objected that such a “solution” is purely verbal, since it requires that a single representation possess what for Kant are the supposedly incompatible properties of being sensible and intellectual.<sup>21</sup> This is to conflate the valid taxonomical point that no representation can be both a concept and an intuition with the dubious claim that, given Kant’s radical separation of the faculties, no representation can combine within itself both sensible and intellectual components. Indeed, the discursivity thesis not only allows for such a combination, it requires it. As Kant puts it at the beginning of the *Analytic*: “It is thus just as necessary to make the mind’s concepts sensible (i.e., to add an object to them in intuition) as to make its intuitions understandable (i.e., to bring them under concepts)” (A51 / B75). The Schematism is an integral part of this project, since it provides the conditions under which the categories can be made sensible, while, at the same time, specifying the results of bringing the form of inner sense under various categories.

We have seen that this combination of properties is possessed by a formal intuition. On the one hand, qua intuition, it is subject to the conditions of sensibility, while, on the other, qua determined by the figurative synthesis, it is subject to the categories. Since transcendental time-determinations possess the

same dual nature, they too should be viewed as formal intuitions. In fact, in both the second and third *Critiques*, Kant explicitly characterizes transcendental schemata in this way.<sup>22</sup>

In an effort to make this somewhat clearer and to gain a more fine-grained understanding of a transcendental time-determination as a formal intuition, let us consider first what it means in general for Kant to “determine an intuition”; then what it means to “determine time”; and, finally, what it means to determine it “transcendentally.” Fortunately, this can be done fairly expeditiously on the basis of what we have already learned.

To begin with, at the empirical level, to determine an intuition is to take it as the perception of a particular something, for example, a house or a certain house.<sup>23</sup> In short, it is to “interpret” it, which, as we have seen, is the work of the imagination and involves subjecting the sensible data to a rule that guides apprehension. As we have also seen, such a rule is not itself a concept, since it is not expressed discursively; though it becomes one when it is so expressed (reflected).

With one significant qualification, a time-determination fits this model. The qualification stems from the purely formal nature of time, which precludes it from becoming itself an object of perception. Consequently, any such determination must occur indirectly, by way of a determination of things in time. In other words, to determine time is to fix the temporal relations of appearances in time: for example, as simultaneous or successive, as *a-b* rather than *b-a*. Although this may likewise be viewed as a matter of rule-guided interpretation, what is thereby determined or interpreted is the relative temporal position of appearances rather than, as in the former case, the sorts of thing they present themselves as being.

Finally, to determine time *transcendentally* is to subject it to an *a priori* rule, that is, a category, which constitutes the “unity” of such a determination. Since the distinctive feature of all categorial determination is its objectifying function, to determine time transcendently is to objectify it. More precisely, it is to represent temporal relations as intersubjectively valid, holding for all subjects with the relevant form of sensibility. Although such determination is more than an interpretation, as a product of imagination rather than of understanding, it is not yet a judgment. Nevertheless, it is the precondition of any judgment about objective temporal relations.

The situation is complicated, however, by the fact (to be explored fully in the next chapter) that time itself cannot be perceived. Accordingly, a transcendental time-determination, like a merely empirical one, functions indirectly

through the determination of the relation of appearances in time. But it differs from the latter in that it involves relating them to what Kant sometimes terms “time in general,” which is the temporalized counterpart of an “object in general.” Just as the latter consists in the unification of representations in accordance with a discursive rule through which they are brought under the objective unity of apperception, so the former is the connection of appearances in time on the basis of the temporal expression or exhibition of such a rule.

Admittedly, this makes transcendental time-determinations formal intuitions in a sense somewhat different from the determinate pure intuitions of space and time. The latter are formal features of space and time themselves, whereas transcendental time-determinations are determinations or necessary characteristics of *things* in time. Nevertheless, this does not preclude them from serving as formal intuitions in a *functional* sense, since like space and time they condition the intuition rather than the thought of objects. Kant underscores this aspect of a transcendental time-determination when he characterizes it as a “formal and pure condition of the sensibility, to which the use of the concept of the understanding is restricted” (3). Nor is this contradicted by the apparently anomalous characterization of a transcendental schema as “the phenomenon, or the sensible concept of an object, in agreement with the category” (8), at least not if, as seems reasonable, one takes “sensible concept” as short for “sensible expression [or exhibition] of a concept.”

Even granting the intuitive nature of transcendental schemata, however, it is often wondered why Kant is preoccupied with time in the Schematism and appears to ignore space. Not only does space supposedly have coequal status as a form of sensibility; but since transcendental time-determinations concern the relations of appearances in (and to) time rather than the nature of time itself, they necessarily involve reference to spatial objects and/or properties. Indeed, Kant himself seems to confirm this in the General Note on the System of Principles added in the second edition, when he remarks that in order to establish the objective reality of the categories, we need not merely intuitions but outer intuitions (B291).<sup>24</sup>

Nevertheless, Kant has good reasons for privileging time and its determinations in the Schematism, and so critics who take him to task for neglecting space are misguided.<sup>25</sup> First, time is characterized in the Transcendental Aesthetic as the “formal *a priori* condition of all appearances in general,” in contrast to space, which is merely the *a priori* condition of outer appearances (A34 / B50). This gives time a greater universality than space, since all appearances, as “modifications of inner sense,” are in time, whereas only outer appearances are

in space. Accordingly, if, as Kant insists, the categories are to apply universally within the field of possible experience, then their application conditions must have reference to time. Second, it is necessary to distinguish between the sphere of objects to which the categories apply and the necessary conditions under which they apply. Kant clearly holds that they apply to objects of both outer and inner sense, but for the reasons cited above, he maintains that they apply to the former in virtue of their temporality. The transcendental synthesis of the imagination is, after all, concerned with the determination of the manifold of inner sense, and thus with the manifold *qua* temporal.<sup>26</sup> Moreover, this essential feature of Kant's position remains unchanged in the second edition.<sup>27</sup>

#### IV. THE PROBLEM OF SCHEMA JUDGMENTS AND THE CONNECTION BETWEEN THE RELATIONAL CATEGORIES AND THEIR SCHEMATA

After characterizing the nature of and the need for transcendental schemata in the first six paragraphs of the Schematism and briefly discussing the schemata of geometrical and empirical concepts in the seventh, Kant turns in the eighth paragraph to the task of specifying the schemata of the various categories. By way of introducing the new topic, he remarks:

Rather than pausing now for a dry and boring analysis of what is required for transcendental schemata of pure concepts of the understanding in general, we would rather present them according to the order of the categories and in connection with these. [A142 / B181]

By describing the task as "dry and boring," Kant seems to attempt to evade the need to provide any argument for the claims connecting the various categories with the specific transcendental time-determinations that serve as their schemata.<sup>28</sup> Instead, as the passage indicates, he contents himself with merely listing the schemata; or, more precisely, he lists the schemata for eight of the twelve categories.<sup>29</sup> This procedure seems a bit cavalier, however, if not utterly question begging. Given the significance Kant attributes to the schemata, it surely is no small matter to determine that a certain schema corresponds uniquely to a particular category. In an effort to address this lacuna in Kant's account, we shall first take up the general issue of the nature and grounds for such a claim of correspondence and then apply these considerations to the connection between the relational categories and their schemata.<sup>30</sup>

### A. The Nature of the Problem

Since the claim that a certain schema corresponds to a particular category is a judgment, it is natural to pose the issue in terms of how such judgments (henceforth referred to as “schema judgments”) should be classified within the Kantian framework. The fact that Kant does little more than list the schemata pertaining to the particular categories might suggest that they are either analytic or merely stipulative. But the heterogeneity of the intellectual and the sensible, taken together with the status of the transcendental schemata as formal intuitions, rules out the former possibility, while the latter is precluded by the fact that it would make the connection between category and schema arbitrary. It is also obvious that a schema judgment cannot be synthetic *a posteriori*, since this implies that the connection between category and schema is based on experience, which is incompatible with the *a priori* status of both. Consequently, it follows by elimination that a schema judgment is synthetic *a priori*.

We need not content ourselves with an argument by elimination, however, since we can also see that these judgments must be classified as synthetic *a priori*. As Lewis White Beck points out, in providing a concept with a schema, “It is not the concept of an intuitive condition which might be added to the concept or included in its definition . . . it is the condition of sensibility itself.”<sup>31</sup> Beck does not characterize such a “condition of sensibility” as a formal intuition, but he does note that it is “a transcendental addendum, a real predicate, a synthetic predicate, a *Bestimmung*, an element in the *ratio essendi* as well as the *ratio cognoscendi*.”<sup>32</sup> Any judgment that provides this is clearly synthetic in Kant’s sense, and if the predicate is an *a priori* representation (as is the case in a schema judgment), then it is also *a priori*.

Nevertheless, such judgments differ significantly from the more familiar synthetic *a priori* judgments based upon them, namely, the principles. Apart from the Postulates, the latter are straightforwardly objective, that is, they involve claims about necessary features of possible experience and the objects thereof. In short, they are “about the world,” not as it is in itself but as it is necessarily experienced by beings with our cognitive faculties. By contrast, a schema judgment is not objective or directly about the world at all. It is rather about the subjective (sensible) conditions under which a pure concept can apply to the world.

Although this may help explain why Kant does not explicitly acknowledge the synthetic *a priori* nature of schema judgments, it does not exempt them from the general requirement of some sort of a deduction. Presumably, this

would consist of two parts: one establishing the general conditions of the possibility of such judgments; the other justifying particular schema judgments. The task of the former might also be expressed as justifying Kant's claim that the "peculiar thing about transcendental philosophy" (the "advantage" over other "didactic sciences" that it shares with mathematics) consists in the fact that, "in addition to the rule (or rather general condition for rules), which is given in the pure concept of the understanding, it can at the same time indicate *a priori* the case to which the rule ought to be applied" (A135 / B174–75). Here Kant is in effect claiming that the categories, like mathematical concepts, carry with them the conditions of their own application. In the case of mathematical concepts this is easily understood, since they are constructible in pure intuition and may even be viewed as rules for such construction. But this is decidedly not the case with the categories, which cannot be so constructed. Moreover, in explaining why transcendental philosophy has this advantage, Kant remarks only that

it deals with concepts that are to be related to their objects *a priori*, hence its objective validity cannot be established *a posteriori*, for that would leave that dignity of theirs entirely untouched; rather it must at the same time offer a general but sufficient characterization of the conditions under which objects in harmony with these concepts can be given, for otherwise they would be without all content, and thus would be mere logical forms and not pure concepts of the understanding. [A135–36 / B175]

The problem with this is that rather than explaining *how* (on what grounds) transcendental philosophy can specify the conditions of the application of its own concepts (the categories), Kant simply asserts that it must be able to do so, if its concepts are to have a real use. But since the latter is just the point at issue, it might seem that this is completely question begging. To take the argument in this way, however, is to ignore the Transcendental Deduction, which has supposedly already shown that the categories are not empty; that they provide conditions of experience and, therefore, have a "transcendental content." More specifically, the second part of the B-Deduction, by connecting the categories with the transcendental synthesis of the imagination, demonstrates that they stand in a necessary connection with time as the form of inner sense. Moreover, this gives us a warrant to suppose that the product of such an imaginative synthesis governed by a specific category will instantiate or "exhibit" what is thought in that category. In short, the general conditions of the possibility of schema judgments are already laid down in the Transcendental Deduction.

The second part of a deduction of schema judgments is a far trickier matter. To begin with, such a deduction cannot itself proceed deductively, that is, it cannot derive schema from category by a purely logical procedure, since that would make the connection analytic. It is rather more like a matter of translating what is thought in a pure concept into temporal terms, into “temporalese,” if you will. But this way of putting it may easily lead to misunderstanding, since it glosses over the crucial point that a thought is being “translated” from its natural discursive form into a non-discursive one.<sup>33</sup> Thus, the model of the translation of a sentence or term from one natural language to another cannot be applied here without significant qualifications.

The same may be said about another superficially attractive model that is sometimes used to interpret Kant’s theory of transcendental schematism, namely, an aesthetic one imported from the third *Critique*.<sup>34</sup> As it is usually understood, this model invites us to view schematizing a concept as a matter of finding in our sensory experience a unity, coherence, or order, which, though itself non-cognitive and non-conceptual, is a necessary condition of the possibility of all rule-governed thought and judgment.<sup>35</sup> Alternatively (on an aesthetic model), it might be seen as a matter of “discovering” a suitable intuitive expression of a category, somewhat like a genius may be said to discover or invent aesthetic ideas that are especially apt for the expression of ideas of reason.<sup>36</sup>

The problem with this model in either form is the precise opposite of the linguistic one. Whereas the weakness of the latter consists in its inability to capture the non-discursive nature of a schema and the role of the imagination in its production, the weakness of the aesthetic model is that it ignores the rule-governed nature of the procedure of determining the schemata of the categories and, therefore, undermines the synthetic *a priori* nature of a schema judgment. Even though Kant notoriously refers to the schematizing process as a “hidden art in the depths of the human soul, whose true operations we can divine from nature and lay unveiled before our eyes only with difficulty” (A141 / B180–81), the determination of the nature of these schemata (as opposed to their production) cannot itself be viewed as such an “art,” at least not if the Kantian project is to make sense.

The inadequacy of these models indicates that what is needed is an understanding of schema judgments that recognizes both the non-discursive nature of a schema and the rule-guided nature of the judgment specifying it. Although this seems like a tall order, it is achievable if we keep in mind that what differentiates the use of the power of judgment in determining the appropriate schemata for the categories from its aesthetic use is that it operates with a de-

terminate criterion of success (a feature that is captured by the linguistic model). Specifically, we know by an analysis of the very terms of the problem that a putative time-determination counts as the schema of a category just in case it makes possible the application to appearances of the category in question, that is, is capable of serving as the “condition of the rule.”

Accordingly, there are two questions to ask of any schema candidate: (1) whether it may be viewed as a transcendental time-determination (by the criteria laid out in the previous section) and (2) whether it is uniquely qualified to serve as the condition of the rule expressed in the category. Since a thorough discussion of this issue would require addressing these questions to each of the schemata, it must here suffice to consider the connection between the relational categories and their schemata.<sup>37</sup> Not only are these schemata more plausibly viewed as transcendental time-determinations than the others, they underlie the Analogies of Experience, which are the concern of the next chapter.<sup>38</sup>

## **B. The Relational Categories and Their Schemata**

According to Kant, the schemata of the relational categories concern the order of time (A145 / B185), which means that they serve to order appearances in time.<sup>39</sup> As such, they clearly fit the characterization of a transcendental time-determination. Thus, the operative question in each case is whether the schema is uniquely qualified to function as the sensible condition that makes possible the application to appearances of the rule contained in the corresponding pure concept. The first of these relational categories is substance, the pure concept of which, as we have already seen, is the concept of something that must always be considered as subject and never as predicate of something else. Accordingly, its schema must allow for the possibility of saying of anything temporal that it is such a subject. Since Kant identifies this schema with the “persistence [*Beharrlichkeit*]<sup>40</sup> of the real in time” (A144 / B183), our first task is to determine whether this meets the above criterion.

The analysis here parallels the previous analysis of the pure concept (chapter 6). We begin by asking for the necessary condition under which we can say of something temporal that it is a real subject or owner of properties rather than a merely logical subject of predicates. The most obvious candidate for such a necessary condition is reidentifiability: only something that is reidentifiable throughout a change of states can be distinguished from one or more of these states and considered to be their “real subject,” that is, as something to which these states pertain as modifications or in which they “inhere” as accidents. But



in order to be reidentifiable the subject must persist through a given period of time. Consequently, at least a relative persistence is a necessary condition for something temporal that is to function as a “real subject.”

Nevertheless, this is not sufficient, since Kant takes the schema of substance to require persistence through *all* time (sempiternity) rather than merely through a certain stretch of time. Thus, the problem is to show that the latter is required. This can easily be done, however, if we recall the distinction drawn in chapter 6 between the strictly judgmental concept of something that is fixed as the subject of a given judgment and the ontological concept of substance, which is the concept of something that for *every* judgmental context must be conceived of always as subject and never as predicate or property of anything else. The basic point is that absolute persistence is required for the application of the latter (the ontological concept) to something in time, but not of the former.

The argument here is straightforward, consisting of nothing more than the extension of the line of reasoning concerning reidentifiability. Thus, just as persistence throughout a certain stretch of time is a condition that must be met by anything temporal that is to serve as a real subject to which properties are attached, so persistence throughout all time is a condition that must be met by anything temporal that is *always* to be conceived of as subject and never as property of anything else. In other words, the concept of something existing in time that must always be regarded as subject and never as property or state is equivalent to the concept of something that is reidentifiable throughout all change. But to be reidentifiable throughout all change is to persist. Consequently, as expressing the condition under which alone the category can be applied to appearances, persistence is the schema of substance. As Kant puts the matter later in the *Critique*:

If I leave out persistence (which is existence at all times), then nothing is left in my concept of substance except the logical representation of the subject, which I try to realize by representing to myself something that can occur solely as subject (without being a predicate of anything). But then it is not only the case that I do not even know of any conditions under which this logical preeminence can be attributed to any sort of thing, it is also the case that absolutely nothing further is to be made of it, and not even the least consequence is to be drawn from it, because by its means no object whatever of the use of this concept is determined, and one therefore does not even know whether the latter means anything at all. [A242–43 / B300–01]

Kant characterizes the pure concept of a cause simply as “something that allows an inference to the existence of something else” (A243 / B301). Since this

inference is based on the thought of the dependence of the latter on the former, the pure relational category of cause and effect is just the ontological counterpart of the logical relation of ground and consequent expressed in the hypothetical form of judgment. More precisely, it is the relation of a real ground to its consequent.

Correlatively, the schema of causality is defined as “the real upon which, whenever it is posited, something else always follows.” Kant further states that it “consists in the succession of the manifold insofar as it is subject to a rule” (A144 / B187). By distinguishing between the definition or characterization of the schema and the statement of what it consists in (which does not occur in any of the other schemata), Kant is presumably attempting to contrast the schema of the concept of a cause as such with that of a causal relation. It is, however, really the latter, which is just rule-governed succession, that is our concern.

Accordingly, the question is whether the notion of a rule-governed succession provides the application condition to appearances given under the form of time of the relation of a real ground to its consequent. Once again, however, this should not be too difficult to answer. The essential point is that the category expresses an ordering rule, specifically an ordering of existential dependence. Consequently, the schema must provide the representation of a temporal sequence exhibiting such dependence. But to represent a sequence of states of affairs or events in time as exhibiting this dependence relation is simply to think of it as governed by a rule of the form: if *A* at *t*<sub>1</sub> then *B* at *t*<sub>2</sub>. Rule-governed succession is, therefore, the schema of the pure concept of causality. It is the condition under which alone the pure concept of existential dependence has any applicability to the data of human experience or, equivalently, it is the “condition of the rule.”

Finally, according to Kant, “The schema of community (reciprocity), or of the reciprocal causality of substances with regard to their accidents, is the co-existence [*Zugleichsein*]<sup>41</sup> of the determinations of the one with those of the other in accordance with a general rule” (A144 / B183–84). Thus, we must see whether similar considerations apply in this case as well. Although there is an extra level of complexity resulting from the obscurity regarding the connection of the category to the disjunctive form of judgment, we shall see that a case can be made here as well.

As was suggested in chapter 6, the key to understanding the connection between the category of community and the disjunctive form of judgment is that the latter embodies the relation of logical coordination (as contrasted with the

relation of subordination embodied in the hypothetical form). Thus, the category expresses the thought of a *real* coordination (between existents) or, more precisely, a thoroughgoing interdependence. Once again, then, what the schema must provide is the temporal expression or exhibition of this relation. But there are only two possible relations of things to one another in time: coexistence and succession. Moreover, the latter is clearly not suitable to the task, since, as we have just seen, rule-governed succession is the schema of the logical subordination or dependence thought in the ground-consequent relation. Thus, if the relation of coordination or interdependence thought in the category of community is to have any temporal expression (as it must if it is to be applied to appearances), then it must involve the relation of coexistence. Clearly, however, mere coexistence will not suffice, since that is a contingent matter and, as such, cannot represent the necessity thought in the category. Thus, *faut de mieux*, it must be a rule-governed coexistence, understood as the coexistence of the changes of one substance with those of others.

## V. THE SCHEMATA AND THE PRINCIPLES

Although we shall deal in the next chapter with the Analogies of Experience, it will be useful to conclude the present one with some general considerations concerning the connection between the schemata and the principles that make use of them. The key to this connection lies in the fact that the transcendental schemata are not only the sensible conditions that both realize and restrict the categories, they are also conditions of the determination of appearances in time, and thus of the possibility of experience. In fact, it is precisely because of this dual function as conditions that they can be said to mediate between the pure concepts and appearances.

The first sense of 'condition' is the main focus of the Schematism and has already been dealt with sufficiently. It is the second sense, however, that is operative in the Principles of Pure Understanding, and it has not yet been considered. Here the main point is that, apart from the Postulates of Empirical Thought, each of these principles is a synthetic *a priori* judgment, which asserts that a particular schema functions as a necessary condition of the possibility of experience.<sup>42</sup> For example, the Axioms of Intuition and the Anticipations of Perception (which will not be further analyzed here) assert, respectively, that everything intuited has an extensive magnitude and is therefore numerable (the schema of quantity), and that every sensation has an intensive magnitude, that is, a degree (the schema of quality). In the case of the Analogies of Experience,

we shall see that each claims that the schema of one of the relational categories functions as a condition of empirical time-determination.

Kant complicates the situation, however, by using the term *analogy* to characterize both the set of principles designated by that name and the general relation between category and schema that applies to all of the principles. In the first case 'analogy' is taken in a mathematical sense as equivalent to 'ratio' or 'proportion'. Kant justifies this choice of terms on the grounds that the schemata involved in these principles correspond to the relational categories (each of which expresses a two-term relation) and that the specific function of these principles is to determine the relation of appearances to one another in a single time. The analogy is thus between the two-term relation expressed in the category and its schema, on the one hand, and the presumed relation of a given appearance to some unspecified relatum, on the other. For example, in the case of the causal relation, the analogy enables us to determine *a priori* that for any given event *y*, there must be some antecedent event *x* from which *y* follows in accordance with a rule. In Kant's own terms, the relation provides "a rule for seeking it [the fourth member] in experience, and a mark for discovering it there" (A180 / B222). Kant notes that the fact that the analogy provides only a rule or decision procedure for seeking the fourth member, rather than the fourth member itself, differentiates "analogies" in philosophy from those in mathematics (A179–80 / B222). It is also the reason why he states that the Analogies (unlike the Mathematical Principles) are valid of appearances only regulatively rather than constitutively (A180 / B223).

In contrast to this reasonably well developed explanation, the second use of 'analogy' seems like a mere afterthought on Kant's part. Nevertheless it is essential to understanding the synthetic *a priori* character of the principles as a whole. This usage occurs in the conclusion to the preliminary discussion of the Analogies of Experience, where Kant writes:

These principles, therefore, justify us in compounding appearances only in accord with an analogy with the logical and general unity of concepts, and hence in the principle itself we make use of the category, but in its execution (its application to appearances) we set its schema in its place as the key to its use, or rather set the latter alongside the former, as its restricting condition, under the name of its formula. [A181 / B224]

Although the location of this passage in the text suggests that Kant is still talking merely about the Analogies, the claim applies to all of the principles.<sup>43</sup>

As the passage indicates, the basic analogy Kant has in mind is between the pure concepts and their schemata, which is based on the view of a schema as providing a temporal expression of what is thought in the concept or category. The ensuing analogy between category and principle is thus attributable to the fact that all the principles make use of the schemata by subsuming appearances under them. This is precisely what makes them both synthetic and *a priori*.

To begin with, the importance of there being an analogy between categories and principles (based on the analogy between the former and their schemata) is that it ensures that the latter contain a categorial element in virtue of which they can function as universal and necessary rules for the unification of appearances. The denial of this analogy is, therefore, tantamount to a denial of the apriority of the principles.

It is equally important, however, that category and schema be *merely* analogous rather than identical. To assert the latter would be to deny the transcendental distinction between the sensible and the intellectual, which would lead, in turn, to the denial of the syntheticity of the principles. Since they would be based on nothing more than an analysis of what is necessary for the unity of thought, they (like the principle of the synthetic unity of apperception) would be analytic. Moreover, if such principles were thought to have reference to objects at all, the latter would have to be defined in purely conceptual terms as "objects of a mere understanding," without any reference to the conditions of sensible intuition. In short, the characterization of the relation between category and schema as one of "analogy" rather than identity is crucial for the syntheticity of the principles as well as for the limitation of their scope to phenomena.

Thus, the presence of schemata accounts for both the apriority and the syntheticity of *all* the principles, not merely the Analogies. Unlike judgments that predicate pure concepts of "objects in general," those that subsume appearances under schemata cannot be regarded as even "covertly analytic," because the formal conditions of sensibility (schemata) under which objects (appearances) are subsumed in these judgments are not themselves contained in the mere concept of an object. Nor can it be argued that these judgments become analytic as soon as we characterize the objects being subsumed under the schemata as themselves temporal. For the determination of the universal and necessary properties of objects qua temporal is itself only possible through synthetic judgments, which link such objects with their temporal conditions.<sup>44</sup> Indeed, the principles are precisely such judgments. But we have seen that the

possibility of these judgments, and, therefore, the possibility of a “metaphysic of experience,” rests upon the possibility of specifying the temporal “analogues” of the categorial rules provided by the pure concepts of the understanding. Moreover, this explains why, in a letter to Reinhold, Kant suggests that it is in the Schematism that we find the real beginning of the account of the possibility of synthetic *a priori* judgments (Br II: 38; 301).

## Chapter 9 The Analogies of Experience

Although the bulk of critical attention remains focused on the Second Analogy, it has become increasingly recognized that the Analogies can only be properly understood if taken together. Moreover, this clearly reflects Kant's own view, since all three are concerned with a single issue, namely, the conditions of time-determination, and each treats one of these conditions, which Kant connects with distinct "modes of time": duration, succession, and coexistence.<sup>1</sup> But since Kant prefaces his discussion of the specific Analogies with a consideration of the underlying problem and includes an analysis of the latter within the Second Analogy as well, the present chapter will be divided into four parts: the first dealing with Kant's analyses of the general problem and the remainder with his account of the three Analogies.

### I. THE GENERAL PROBLEM OF THE ANALOGIES

We have seen that whereas the "mathematical" principles of extensive and intensive magnitude are concerned with the structure of empiri-

cal intuition or perception, the Analogies, as “dynamical” principles, are concerned with the *experience* of an objective temporal order of appearances. Thus, with the Analogies we move from a consideration of the transcendental conditions of perception, which involve rules of apprehension guiding an imaginative synthesis, to a consideration of the transcendental conditions of the cognition of an objective temporal order of existences corresponding to these perceptions. By way of underscoring this, in the *Prolegomena* Kant urges the reader “to heed well this distinction of experience from a mere aggregate of perceptions and to judge the mode of proof [of the Analogies] from that standpoint” (Pro 4: 310; 103). In the *Critique*, Kant discusses this mode of proof in two places, which form a prelude to the analyses of the arguments of the specific Analogies.

### A. The Transcendental Setting

(A189 / B234–A191 / B236)

Although Kant formulates a general principle of the Analogies and provides a brief argument for it in the first edition, the best introduction to the problematic of the Analogies in the initial version of the *Critique* is located in the first paragraph of the Second Analogy.<sup>2</sup> Since, in spite of this location, it is here that Kant raises the general problem of explaining how cognition of an objective temporal order is possible, it may be taken as an introduction to the Analogies as a whole.<sup>3</sup> The paragraph falls quite neatly into four parts. The first raises the problem of explaining the possibility of cognition of an objective temporal order. The second argues in effect that transcendental realism cannot account for this possibility. The third formulates the problem in terms of the language and assumptions of transcendental idealism. The fourth sketches the “critical” or transcendently idealistic solution.

*I. The Cognition of an Objective Temporal Order.* By an “objective temporal order” is meant simply an order of occurrences in the world. The general problem with which all the Analogies are concerned is the possibility of the cognition of such an order. It is not, however, immediately apparent that there is any problem here, at least not one requiring a transcendental solution. Moreover, Kant’s own explanation serves more to obscure than to clarify the issue. As he puts it,

The apprehension of the manifold of appearance is always successive. The representations of the parts succeed one another. Whether they also succeed in the object is a second point for reflection, which is not contained in the first. [A189 / B234]



This suggests that the problem is that, since apprehension is always successive, an inspection of the order of apprehension, that is, the order in which representations occur in consciousness, does not provide adequate evidence for forming reliable judgments about the quite different order of what is represented. Thus, from the succession of my representations *a-b*, I cannot infer anything about the order of the object represented. To cite Kant's example: I apprehend the parts of a house successively, yet I judge them to be coexisting parts of an enduring object. How is such a judgment possible?

It is frequently claimed that Kant finds a problem here only because of a dubious psychological assumption that all apprehension is in fact successive. In response it is sometimes argued that Kant does not really need this assumption, for all that is required to generate the problem is the non-controversial claim that the order of apprehension is not a reliable indicator of the objective order. The two orders may, but need not, coincide.<sup>4</sup> Although this response is correct as far as it goes, it shares with the original objection the erroneous assumption (admittedly suggested by Kant's formulation) that the problem lies in having to make a judgment about the objective order on the basis of the subjective one.

In order to understand the problem that concerns Kant in the Analogies, it is necessary to reject the latter assumption. The subjective order is not a datum on the basis of which the mind must somehow infer or construct an objective order. It is rather what would remain if (*per impossibile*) we could remove the determinate structure imposed on the sensibly given (the manifold of inner sense) by the understanding. Thus, what Kant is trying to say here is that, if all we had were this indeterminate subjective order, we would not be able to represent any temporal order at all (objective or subjective).<sup>5</sup> The problem, then, is to explain how time consciousness, that is, cognition of a temporal order, is possible. In Kant's terms, it is to provide the "formal conditions of empirical truth." Not only is this a significant problem, it is also the very problem that was left unresolved by the Transcendental Deduction.

2. *The Inadequacy of Transcendental Realism.* It has been suggested that the second portion of the paragraph be taken to be arguing that transcendental realism is incapable of accounting for the possibility of cognition of an objective temporal order. Since here Kant does not explicitly refer to transcendental realism, this may seem to be a fanciful reconstruction. Nevertheless, it is not without a basis in the text. Consider Kant's characterization of the problem in terms of his transcendental distinction:

If appearances were things in themselves, then no human being would be able to assess from the succession of representations how the manifold is combined in the object. For we have to do only with our representations; how things in themselves may be (without regard to representations through which they affect us) is entirely beyond our cognitive sphere. [A190 / B235]

Here Kant is maintaining that, if empirical objects are regarded as things in themselves rather than as appearances, it is impossible to understand how we could have any knowledge of their objective temporal order. This is because “we have to do only with our representations,” whereas the order in question is, by definition, distinct from the order in which the representations occur in consciousness. Since transcendental realism regards what for Kant are “mere appearances” as things in themselves, it cannot account for the possibility of cognition of such an order.

This is nothing more than an application of the general point made in the A-version of the Fourth Paralogism to the specific problem of the cognition of a temporal order: “If we let outer objects count as things in themselves, then it is absolutely impossible to comprehend how we are to acquire cognition of their reality outside us, since we base this merely on the representation which is in us” (A378). Consequently, the argument certainly applies to empirical idealism, that is, the Cartesian-Lockean version of transcendental realism that is under attack in the Paralogisms. One might ask, however, whether it can be made applicable to *all* forms of transcendental realism, including those that for one reason or another might reject empirical idealism.

We see that it can when we consider Kant’s claim that “time in itself” or “absolute time” cannot be perceived, which is implicit in the present argument. This unperceivability precludes our directly comparing our representations with a pre-given temporal order that is assumed to be transcendently real. Presumably, this would be granted by all forms of transcendental realism, including Newton’s.<sup>6</sup> But if “time itself” cannot be perceived, it follows that the only way to determine an objective temporal order is through the ordering of “appearances” in it. At this point, however, the transcendental realist must admit that the only order that is actually “given” to the mind is that of the occurrence of its own representations. The latter, then, will be the only “object” to which the mind has any access on a transcendently realistic account. Thus, the problem of access, which empirical idealism raises about material objects and which leads to a skepticism about the “external world,” recurs in the case of an objective temporal order for transcendental realism. In fact, it is the transcendental realist (not Kant) who is in the impossible situation of having to

make inferences about a putative objective temporal order on the basis of a subjective one.

3. *The Idealistic Reformulation.* Given the way in which the problem of explaining the possibility of cognition of an objective temporal order has been presented, it is by no means obvious that the transcendental idealist is in any better position to resolve it than the transcendental realist. Thus, while the latter has the problem of explaining the possibility of access to an objective temporal order of things in themselves, the former has the problem of explaining the possibility of distinguishing between an objective and a subjective temporal order *within* the realm of appearance.

After introducing the example of the perception of a house, which is designed to illustrate the point that we do not identify the order of the representing of the parts with an order of successive states in the object represented, Kant poses the problem in explicit terms. “[A]s soon,” he writes, “as I raise my concept of an object to transcendental significance, the house is not a thing in itself at all but only an appearance, i.e., a representation, the transcendental object of which is unknown” (A190–91 / B235–36). And this, in turn, leads him to ask:

[W]hat do I understand by the question: how the manifold may be combined in the appearance itself (which yet is nothing in itself)? Here that which lies in the successive apprehension is considered as representation, but the appearance that is given to me, in spite of the fact that it is nothing more than a sum of these representations, is considered as their object, with which my concept, which I draw from the representations of apprehension, is to agree. [A191 / B236]

The key to understanding this difficult passage lies in recognizing the progressive, dialectical nature of Kant’s account. He begins by raising the question of the conditions of the possibility of making judgments about an objective temporal order, which is relevant to both the transcendental realist and the transcendental idealist, though the former is incapable of answering it. But the transcendentially idealistic analysis of the problem leads to a new question about the very concept of an objective temporal order of appearances. This question is unique to the transcendental idealist.

Further reflection, however, suggests that these two questions are really equivalent. As we have already seen, it is the very essence of Kant’s “transcendental turn” that the meaning of ‘object’ must be explicated in terms of the conditions of the representation of objects. All that Kant is doing here is applying this principle to a special sense of ‘object’; namely, an objective temporal order.

Since it is only in and through judgment that we can represent objects, determining the conditions or grounds for making judgments about an objective temporal order also explains what is represented in the thought of such an order. In other words, the answer to the first of these questions will also be an answer to the second.

4. *The “Critical” Solution.* In the remainder of the paragraph Kant gives the gist of his solution to the problem. He thus explains how, through successive representations, we can represent to ourselves an object distinct from our representations, even though the object is nothing but the sum of these representations. Kant’s answer is that we represent appearance as an object, in this case an objective temporal order, by subjecting our representations to a rule. Correlatively, the “object” here is just the temporal order of given appearances that is thought as the result of the subjection of the representations to a rule. As Kant puts it at the conclusion of the paragraph:

[A]pppearance, in contradistinction to the representations of apprehension, can thereby only be represented as an object that is distinct from them if it stands under a rule that distinguishes it from every other apprehension and makes one way of combining the manifold necessary. That in the appearance which contains the condition of this necessary rule of apprehension is the object. [A191 / B236]

Insofar as it suggests that the rule in question is a “rule of apprehension,” this passage is a bit misleading. As we shall see in more detail in the sequel, it is rather a rule for conceptualization or judgment, which dictates how the given representations are to be connected in an objective unity distinct from the connection they contingently have in apprehension.<sup>7</sup> Correlatively, the “necessity” here consists in the conceptual constraint on the thought of the temporal order, through which the latter acquires objective validity. Finally, as always for Kant, this necessity, and with it the objective validity of the thought, is produced by the imposition of an *a priori* rule. Accordingly, the three Analogies will be concerned with the distinct rules required for the cognition (experience) of the objective relations of appearances in time.

#### B. The General Principle of the Analogies in B (B218–19)

As formulated in the second edition, the general principle reads: “**Experience is possible only through the representation of a necessary connection of perceptions**” (B218). The contrast between experience and perception indi-

cates the close connection between this principle and §26, and the argument for it consists of five steps:

[Step 1.] Experience is empirical cognition, i.e., a cognition that determines an object through perceptions. It is therefore a synthesis of perceptions, which is not itself contained in perception, but contains the synthetic unity of the manifold of perception in one consciousness, which constitutes what is essential in a cognition of **objects** of the senses, i.e., of experience (not merely the intuition or sensation of the senses).

We have seen that the main thrust of §26 is that perception itself requires a synthesis, the implications of which are explored in detail in the Mathematical Principles. Now, developing the point merely stated in passing in §26, Kant indicates that a whole new level of synthesis, the materials for which are provided by the already synthesized perceptions, is required for experience, that is, for empirical cognition of the objects perceived.

[Step 2.] Now in experience, to be sure, perceptions come together only contingently, so that no necessity of their connection is or can become evident in the perceptions themselves, since apprehension is only a juxtaposition [*Zusammenstellung*] of the manifold of empirical intuition, but no representation of the necessity of the combined existence of the appearances that it juxtaposes in space and time is to be encountered.

By “in experience” Kant here obviously means in perception or apprehension. Thus, his point is that the way in which the manifold of empirical intuition is “juxtaposed” in perception is “contingent” in the sense that it is based on factors relative to the situation of a perceiver, which, therefore, cannot of itself license an objectivity claim. Once again, the latter always requires a rule-governed connection, which introduces “necessity” in the form of a conceptual constraint. What Kant adds here to this familiar story is a reference to the “existence of appearances” and their juxtaposition in space and time. Since by the former Kant means their spatiotemporal existence, that is, their determinate location and relations in space and time, these are intimately related.

[Step 3.] But since experience is a cognition of objects through perception, consequently the relation in the existence of the manifold is to be represented in it not as it is juxtaposed in time but as it is objectively in time. . . .

Here Kant contrasts the objective temporal order of existence with the contingent, subjective order of apprehension. Since it is the former that is experienced in the proper sense, it is the object of concern.

[Step 4.] [Y]et since time itself cannot be perceived, the determination of the existence of objects in time can only come about through their combination in time in general, hence only through *a priori* connecting concepts.

This is the crucial step since it defines the problem for all three Analogies. As I have already indicated, the thesis of the unperceivability of time entails that perceptions do not appear with their temporal order, as it were, stamped on them, though they do present themselves as extensive and intensive magnitudes. Correlatively, “time in general” functions as the temporalized form of “object in general.” Just as to relate one’s representations to the latter is to unify them in accordance with *a priori* concepts that ground the objective validity of the synthesis, so to relate appearances to the former is to unify them according to rules that determine their temporal order for all subjects of experience (“consciousness in general”). In other words, it is by subjecting appearances to such rules that one represents to oneself an objective temporal order.

[Step 5.] Now since these always carry necessity along with them, experience is thus possible only through a representation of the necessary connection of the perceptions.

Here Kant makes explicit the conclusion arrived at in the preceding step and already foreshadowed in §26. Assuming, as seems reasonable, that experience involves the cognition of an objective temporal order of appearances, it follows that the rules required for the cognition of such an order are necessary conditions of experience. Accordingly, the task of the Analogies is to specify these rules and to demonstrate their necessity.

## II. THE FIRST ANALOGY

The First Analogy is concerned with the applicability of the schema of substance: the persistence of the real in time. Although Kant suggests at one point (A215 / B262) that the function of this Analogy is to serve as a condition of the determination of duration, that is, of the measurement of time, it is really concerned with the necessary conditions of all time-determination.<sup>8</sup> What Kant must show is that only some thing or things that persist throughout all time can provide the requisite condition for the unification of things and events in a single time, and thus in a single experience.<sup>9</sup>

The situation is complicated, however, by Kant’s combining this general thesis about the necessity of something persisting as a condition for all time-deter-

mination with a thesis about change. He argues that all “change” (*Wechsel*) among appearances must be conceived and experienced as an alteration (*Veränderung*) of a substance that persists. Thus, in the first edition the principle states: “All appearances contain that which persists (**substance**) as the object itself, and that which can change as its mere determination, i.e., a way in which the object exists” (A182). In the second edition Kant makes what appears to be an even stronger claim, namely, “In all change of appearances substance persists, and its quantum is neither increased nor diminished in nature” (B224). By thus claiming that the quantum of substance remains constant, Kant has been thought by many critics to have been engaged in the disreputable project of attempting to provide a transcendental proof of an empirical thesis.

These, then, constitute the basic issues we must deal with. In particular, by focusing on the argument added in the second edition, the aim is to show that the argument for absolute persistence is sound and that the persistence of the quantum of substance that it affirms must be distinguished from the principle of the conservation of matter, for which Kant argues in the *Metaphysical Foundations of Natural Science*.

The main argument, which is contained in the first paragraph in the second edition (B224–25), can be broken down into seven steps, which can, in turn, be grouped into four parts or subarguments. The first part (steps 1–4) contends that something at least relatively persisting is required as a substratum or backdrop in relation to which change can be experienced. The second part (step 5) argues that every change of appearances must be regarded as the change of state of this substratum. The third part (step 6) asserts that this substratum must be absolutely, not simply relatively, persistent. The final part (step 7) maintains that the quantity of this persisting substratum remains constant throughout all change. It should be clear from this outline that the argument has a progressive structure, each step presupposing and building upon the preceding.

### A. The Backdrop Thesis

[Step 1.] All appearances are in time, in which, as substratum (as persisting form of inner intuition), both **simultaneity** as well as **succession** can alone be represented.

Here Kant reaffirms the temporality of all appearances, thereby reiterating the contention of the Transcendental Aesthetic that “[t]ime is the formal *a priori* condition of all appearances in general” and, therefore, that “all appearances in general, i.e., all objects of the senses, are in time, and necessarily stand in relations of time” (A34 / B50–51). The foundational role of time is now expressed

by characterizing it as 'substratum', which is a key term in the overall argument, since it prepares the way for the eventual introduction of the concept of substance. At this point, however, all that Kant can legitimately claim is that the representation of time must be presupposed in order to represent either simultaneity or succession.

[Step 2.] The time, therefore, in which all change of appearances is to be thought, lasts and does not change; since it is that in which succession or simultaneity can be represented only as determinations of it.

Against this, it is frequently argued that, though it is true that time does not change, but rather that things change in time, it is equally true that time does not remain or endure. As Edward Caird put it:

It may be objected that to say that "time itself does not change" is like saying that passing away does not itself pass away. So far the endurance of time and the permanence of the changing might even seem to mean only that the moments of time never cease to pass away, and the changing never ceases to change. A perpetual flux would therefore sufficiently "represent" all the permanence that is in time.<sup>10</sup>

Caird's contention is true but irrelevant. The essential point is that the constant flux occurs in a single time.<sup>11</sup> The claim that time is unchangeable is equivalent to the claim that it retains its identity as one and the same time (temporal framework) throughout all change. The most that Kant can be charged with here is a lack of clarity, though it is difficult to imagine what else he could have meant. Moreover, we shall see that this is precisely the sense in which substance is said to be unchangeable or persistent. That is also why the thesis of the unity of time is central to the whole argument, by making it possible to link substance directly to time and to argue that the schema of substance is necessary for its determinate representation.

[Step 3.] Now time cannot be perceived by itself.

As already noted, this is an essential step in the argument of each of the Analogies. Its significance stems from the fact that it at once precludes the possibility of determining the temporal order of appearances by referring them to time itself as a quasi-object and requires that this order be determined "immanently" by connecting these appearances according to rules for their unification in consciousness.

[Step 4.] Consequently it is in the objects of perception, i.e., the appearances, that the substratum must be encountered that represents time in general and in which all



change or simultaneity can be perceived in apprehension through the relation of the appearances to it.

This is the first step in the “immanentization” of the temporal order of appearances that is necessitated by the unperceivability of time. Given this unperceivability, it becomes necessary to presuppose some perceptually accessible model for time itself as a condition of the possibility of determining the temporal relations of appearances. In the first edition, Kant identifies this model or, as he terms it, “substratum,” with the “object itself” (A182–83). The basic point is that the latter must somehow embody the unchangeableness or persistence that has already been attributed to time itself. If there were nothing that persists, if everything were in constant flux, then we could not even be aware of succession as such, not to mention simultaneity. Consequently, an enduring, perceivable object (or objects) is required to provide the backdrop or frame of reference by means of which the succession, simultaneity, and duration of appearances in a common time can be determined.

This will be termed the “backdrop thesis,” and it is the result of the first four steps of the argument. Although itself not a trivial result, it is far from the overall conclusion at which the argument aims. Thus, the major interest of the First Analogy centers around Kant’s efforts to go beyond this result in the next three parts of the argument.

## B. From Substratum to Subject

[Step 5.] However, the substratum of everything real, i.e., everything that belongs to the existence of things, is **substance**, of which everything that belongs to existence can be thought only as a determination.

Although the expression “everything that belongs to the existence of things” is highly obscure, it seems reasonable to follow Paton in taking it to refer to changing appearances having determinate positions in time.<sup>12</sup> On this reading, then, Kant is claiming that all these appearances must be regarded as states or determinations of substance. At this stage of the argument there is certainly no warrant for taking ‘substance’ in the full-blown ontological sense, as referring to something absolutely persistent. But even if we understand it to refer merely to relatively persisting phenomenal entities, the claim still takes us considerably beyond the backdrop thesis. Whereas that asserts only that the presence in experience of some relatively persisting entity or entities is a necessary condition of the possibility of the experience of the succession or simultaneity of appearances in time, the present claim is that all changes (*Wechseln*) of appearances

(where one appearance is succeeded by another) must be experienced as alterations (*Veränderungen*) in the states of these entities.

Kant's most explicit formulation of this claim is at the beginning of the Second Analogy, where he remarks that one way of expressing the result of the First Analogy is in terms of the principle "**All change (succession) of appearance is only alteration**" (B233). In Van Cleve's more precise reformulation, "For any  $x$ , if  $x$  changes, there is a  $y$  such that  $y$  alters in respect of  $x$ ."<sup>13</sup> Admittedly, Kant misdescribes the situation by suggesting that this principle is itself the final conclusion of the argument; nevertheless, it is certainly an essential step (perhaps *the* essential step). Our present concern, therefore, is to see if it is possible to find an argument in support of it.

As a first step, it is crucial to become clear about the meaning of '*Wechsel*', which both Kemp Smith and Guyer and Wood usually translate simply as 'change'. Since Kant frequently uses the term to refer to an "arising" (*Entstehen*) or a "perishing" (*Vergehen*), it might be thought that Kant means by it something like the Aristotelian notion of substantial change. But it would be misleading to render '*Wechsel*' in this way, since Kant's main thesis is that only states or determinations of substances *wechseln*, not the substances themselves. His own example of *ein Wechsel* is the change that occurs when a piece of wood is burned. In ordinary parlance we might say that the wood "becomes" smoke and ashes, just as we might say of a successful experiment in alchemy that the base metal "becomes" or is "changed into" gold. But since neither the wood nor the base metal itself survives the process, what is really meant is that what initially appeared in the form of wood or base metal (the matter) is transformed in the process into smoke and ashes, in the one case, and gold, in the other. Although this is precisely what the argument must show, for the present the main point is simply that *ein Wechsel* is the kind of change in which one item is replaced by another. Accordingly, it will here be characterized as a 'replacement change'.<sup>14</sup>

Expressed schematically, what must be shown is that every replacement of a given state of affairs ( $x$ ) at  $t_1$ , by some contrary state of affairs (non- $x$ ) at  $t_2$ , must be conceived and experienced as the alteration (change of state) of some entity ( $y$ ) that endures throughout the process. Correlatively, as the combustion and alchemy examples indicate,  $x$  and non- $x$  must be thought as successive determinations of  $y$ . Unfortunately, since Kant is ultimately concerned in the First Analogy with the demonstration of the necessity of some thing or things that are absolutely persistent or sempiternal, it is difficult to locate an argument

devoted specifically to this subordinate but essential thesis. Perhaps the closest thing to it is contained in the following passage:

Alteration can therefore be perceived only in substances, and arising or perishing per se cannot be a possible perception unless it concerns merely a determination of that which persists, for it is this very thing that persists that makes possible the representation of the transition from one state into another, and from non-being into being, which can therefore be empirically cognized only as changing determinations of that which lasts. [A188 / B231]

The references to the representation of the transition from one state to another and to empirical cognition indicate that the required argument turns on the conditions of the possibility of the cognition of such a change. Although it may not be immediately evident that any such argument is contained in the above-cited passage, it seems possible to find at least the sketch of one.<sup>15</sup>

Like the argument for the backdrop thesis, this argument requires the premise that time cannot be perceived. Consequently, a single observation is never adequate to determine that a change of any sort has taken place, and *a fortiori* not a replacement change. Any such experience requires two successive observations and the noticing of some difference between what is observed in each case. One cannot, however, infer from a difference in two successive observations that a replacement change has in fact occurred. From all that can be determined from the two observations alone, one could simply be having successive observations of coexisting states of affairs.

For example, I perceive my desk at  $t_1$  and my bookcase at  $t_2$ , but I do not infer from the succession of perceptions that a replacement change has occurred, that is, I do not assume that the desk has somehow “become” or been changed into the bookcase. Again, let us suppose that during the interval between  $t_1$  and  $t_2$  the desk was removed and the bookcase put in its place. I would certainly experience this as a change, but not as a replacement change (at least not in the sense in which the term is being used here). If, by contrast, I experience or believe that I experience a genuine replacement change (as in the cases of combustion), then I am constrained to refer the successive states of affairs to some common subject and to view this occurrence as an alteration in its states. Only by so doing can I represent through my successive perceptions or observations the replacement of one state of affairs by another.<sup>16</sup>

The crucial point here, which admittedly is somewhat obscured by the empirical examples and the talk about observations, is that the assignment of the

successively represented states of affairs to an enduring substratum (as its successive states) functions as the rule through which we think such a change. It can also be described as the form of the thought of a replacement change in the sense that to think such a change (as an object of possible experience) is just to connect one's perceptions according to the rule.

### C. From Relative to Absolute Persistence

[Step 6.] Consequently that which persists, in relation to which alone all temporal relations of appearances can be determined, is substance in the appearance, i.e., the real in the appearance, which as the substratum of all change always remains the same.

Up to this point, the argument has shown only that one must include in one's ontology enduring, reidentifiable entities that function as substrata of change. The next and decisive step requires demonstrating that some of these entities, or perhaps something more fundamental, persist throughout all time. Only by so doing can Kant establish the objective reality of the schema of the pure concept of substance. It is precisely at this point, however, that virtually all of the commentators demur on the grounds that Kant's argument is at best only capable of proving the need for relative persistence and that the move to the absolute variety is totally unwarranted.<sup>17</sup>

At first glance at least, this criticism seems well founded, since there are places where Kant appears to assume that anything that functions substantively in experience, that is, anything that serves as a substratum of change, must for that very reason be sempiternal.<sup>18</sup> Nevertheless, since Kant was undoubtedly aware that the sorts of things that normally function in this way, such as tables, trees, mountains, and planets, themselves come into and go out of existence, it seems more reasonable to construe him to be claiming that such changes must be experienced as alterations of something truly substantial that persists throughout all change.

The argument required at this point turns on the necessary unity or identity of time as a condition of the unity of experience. It proceeds by applying the principle that all replacement change is alteration to the enduring entities or substance candidates whose necessity has been established in the preceding step. Thus, far from assuming that these are sempiternal, it assumes that they are not and considers the conditions of the possibility of the experience of their arising or perishing. Since any such occurrence involves the replacement of one state of affairs by its contrary, it would count as a replacement change. We have already seen, however, that the experience of any such change requires that

both states of affairs be linked to an identical subject ("the object itself") as its successive determinations. But an absolute arising or perishing would be an occurrence in which, *ex hypothesi*, these conditions do not obtain. Moreover, since there would be no way in which the emergence of this new state of affairs could be connected empirically with the preceding time (in which the contrary state of affairs existed), such an occurrence would cause a rupture in the unity of time, and therefore in the unity of experience.<sup>19</sup> Kant sums up the matter at the end of the Analogy:

Substances (in appearance) are the substrata of all time-determinations. The arising of some of them and the perishing of others would itself remove the sole condition of the empirical unity of time, and the appearances would then be related to two different times, in which existence flowed side by side, which is absurd. For there is only **one** time, in which all different times must not be placed simultaneously but only one after another. [A188–89 / B231–32]

The point can be clarified by a brief consideration of Kant's own well-known illustration of the principle, namely, the Lavoisier-inspired example of the burning of a piece of wood:

A philosopher was asked: How much does the smoke weigh? He replied: If you take away from the weight of the wood that was burnt the weight of the ashes that are left over, you will have the weight of the smoke. He thus assumed as incontrovertible that even in fire the matter (substance) never disappears but rather only suffers an alteration in its form. [A185 / B228]

Clearly, the burnt piece of wood must be assumed to have existed for a period of time prior to its destruction by fire and to have been capable of being altered in any number of ways during that period without losing its identity. Just as clearly, its destruction by fire cannot be regarded as simply another alteration of the wood, since it is no longer identifiable as wood at the end of the process. Nevertheless, we still regard the process of combustion as an alteration. The difference is that, instead of treating the piece of wood as the subject that alters, we presuppose some matter that at one stage of its career assumed the form of wood and at a later stage was transformed into smoke and ashes. In other words, in order to conceive of such a transformation it is necessary to consider the piece of wood as the temporary form, state, or determination of some enduring matter. Correlatively, this matter of which things are composed is the "ultimate subject" of predication or, equivalently, "the substantial" in things. As such, it must be presupposed to endure throughout all change, including the arising and perishing of its particular configurations or determinations.

Although the basic claim goes back at least to Aristotle, Kant breaks with the tradition by giving this principle an epistemic rather than a logical or ontological grounding. Thus, instead of claiming that the thought of something simply arising or perishing is self-contradictory, he maintains that such an “occurrence” (like an “event” without a cause) could not be an object of possible experience. This is because if it were to occur, “everything would disappear that alone can represent the unity of time, namely, the identity of the substratum in which alone all change has its thoroughgoing unity” (A186 / B229). In other words, absent a persisting substratum (or substrata) of change, it would be impossible to take one’s perceptions as perceptions of occurrences in a single time, which, in turn, would preclude their connection in a single experience.

#### D. The Quantum of Substance

[Step 7.] Since this, therefore, cannot change in existence, its quantum in nature can also be neither increased nor diminished. [B224–25]

Here Kant goes beyond what has been established so far by asserting, in effect, that the persistence of substance entails the conservation of its quantity in the universe. In spite of the fact that Kant explicitly makes this claim only in the second edition, the discussion of the philosopher’s answer to the query regarding the weight of smoke indicates that it is already implicit in the first. Consequently, it seems appropriate to regard it as an essential aspect of Kant’s theory of substance, rather than a mere afterthought. At the same time, however, it is also an aspect of the theory that is widely rejected. The basic charge is that it involves an illicit move from transcendental to empirical considerations, in particular, that Kant is attempting to “deduce” the principle of the conservation of mass as it is understood within Newtonian mechanics.<sup>20</sup>

This objection stems, however, from a failure to distinguish the transcendental-level argument of the First Analogy from the argument of the *Metaphysical Foundations of Natural Science*, where Kant does affirm the principle of the conservation of matter by “applying” the transcendental principles of the *Critique* to the empirical concept of matter. In order to underscore the difference, we shall briefly consider each argument in turn.

Although in the *Critique* Kant does not offer an argument at this point, it is easy enough to provide one. We have just seen that substance, the ultimate subject of predication, must be identified with the matter of which things are composed and that this matter must be conceived as permanent (in order to function as such a subject). Thus, the question becomes how we are to characterize

this matter in a transcendental account that does not make use of any empirical assumptions. Moreover, Kant himself answers this question for us in the *Architectonic of Pure Reason*, where he characterizes matter, so conceived, as “impenetrable, lifeless extension” (A848 / B876). The point is that the only property that can be legitimately assigned to matter in a transcendental account is the occupation of space.

Matter so conceived, is completely indeterminate. Unlike Aristotelian prime matter, however, which is indeterminate in the metaphysical sense that it is a pure stuff, literally without properties, this matter is indeterminate in the methodological sense that no properties other than the occupation of space can be legitimately assigned to it in a transcendental account. But, so considered, the only category available for the conceptualization of matter is quantity. Consequently, the persistence of matter must be conceived of as the persistence of its quantity.<sup>21</sup>

In contrast to this conceptual claim, which cannot be equated with either a specific law of nature or a conservation principle, Kant attempts in the *Metaphysical Foundations of Natural Science* to derive a conservation principle, which he terms the First Law of Mechanics, namely, “In all changes of corporeal nature the quantity of matter as a whole remains the same, neither increased nor diminished” (MAN 4: 541; 249). Although Kant characterizes this as an *a priori* law of nature, the main point is that its derivation, as well as the derivations of the other laws considered in that work, requires the introduction of empirical premises. Specifically, it appeals to the concept of matter as the “movable in space.” Movability is, therefore, the “empirical” feature which differentiates this definition from that of the *Critique*, and which supposedly makes possible the derivation of specific laws of nature.<sup>22</sup>

In the case of the principle of the conservation of matter, the key move is the determination of the nature of material substance. Kant attempts to accomplish this by combining his definition of matter with the nominal definition of substance as “the ultimate subject of existence, that is, that which does not belong in turn to the existence of another merely as predicate” (MAN 4: 503; 214). The point is that only matter, defined as the movable in space, satisfies the definition of substance, because apart from matter (so defined) no real subject for the properties or accidents of the objects of outer sense can be thought except for space itself. The Transcendental Aesthetic has shown, however, that space is not itself an object but rather the form or condition of our representation of objects of outer sense. Consequently, the movable in space is the only available candidate for substantival status.

As a direct consequence of this thesis, Kant further contends that the independently movable parts of material substance are themselves substances, and that by the 'quantity of matter' is to be understood the number of such substances of which a given portion of matter is composed. The argument here is that any such particle of matter, insofar as it is capable of moving independently of the other particles, is likewise capable of functioning as a subject to which properties or accidents can be attached (MAN 4: 503; 215). Any such object, therefore, fits the definition of substance.<sup>23</sup>

Given this conception of material substance, all that Kant requires to derive the principle of the conservation of matter is the doctrine of the First Analogy that substances cannot be created or destroyed. Since the quantity of matter is defined in terms of the number of substances (independently movable particles) of which matter is composed, this quantity can be changed only through the addition or subtraction of substances. But this would require the creation or annihilation of substances, and that is ruled out by the First Analogy. As a result, the quantity of matter in nature as a whole must remain constant throughout all time (MAN 4: 541–42; 249–50).<sup>24</sup>

### III. THE SECOND ANALOGY

Once again Kant formulates the Analogy differently in the first and second editions. In the first it is called the "Principle of Generation," and it states: "Everything that happens (begins to be) presupposes something which it follows in accordance with a rule" (A189). In the second it is termed more elaborately the "Principle of temporal sequence according to the law of causality," and it states: "All alterations occur in accordance with the law of the connection of cause and effect" (B232).

The first edition formulation is reminiscent of Hume's characterization in the *Treatise* of the causal principle as the maxim that "*whatever begins to exist must have a cause of its existence*."<sup>25</sup> It thus suggests that the target is Hume's challenge to the general principle of causality in the *Treatise*, rather than his challenge in the *Enquiry* to the arguably different principle that similar causes produce similar effects. Following Lewis White Beck, we shall call the former the "every-event-some-cause" principle and the latter the "same-cause-same-effect" principle.<sup>26</sup>

As is borne out by its opening paragraph, the reformulation in the second edition is obviously intended to bring the argument into closer connection with the First Analogy. Just as the latter argued that all change is merely the al-



teration of substance that persists, so the Second Analogy will argue that all such alterations are governed by the “law of the connection of cause and effect.” But since this law is equivalent to the Principle of Generation, the difference between the two formulations is merely terminological. In both editions the goal is to establish the every-event-some-cause principle.

The basic concern of this section is to analyze and evaluate Kant’s argument in support of this principle, and the discussion is divided into four parts. The first deals with some important preliminaries, including definitions of the key terms. The second analyzes the essential argument contained in both editions. The third discusses some interpretive issues concerning the precise nature of Kant’s claim and attempts to sketch a defense of the argument against some familiar objections. The fourth considers the question of the connection between the transcendental principle of causality and particular causal laws.

### A. Some Preliminaries

Before we consider Kant’s argument, it is important to clarify some terminological and interpretive points, which include the key terms ‘cause’ and ‘event’, and the deeply ambiguous conception of an objective succession. To begin with, Kant agrees with Hume that necessity is the essential as well as the problematic feature of the concept of causality. Thus, he claims that “this concept always requires that something A be of such a kind that something else B follows from it necessarily and in accordance with an absolutely universal rule” (A91 / B124); and again, that “the concept of cause brings the trait of necessity with it, which no experience at all can yield” (A112–13). Although Kant does not explain what he means by ‘necessity’ in this context, his use of the term suggests that here, as in the Introduction to the *Critique* (B4–5), it must be taken as equivalent to ‘strict universality’ or ‘invariability’. To claim that *A* is the cause of *B* is to claim that given an *A*-type event, together with certain standing conditions, a *B*-type event will invariably (*jederzeit*) follow.

The second key term is ‘event’ (*Begebenheit*, *Ereignis*, *Wirklichkeit*), which Kant generally treats as synonymous with ‘happening’ or ‘occurrence’ (*Geschehen*). This is potentially a source of some confusion, but the main point is that all of these terms refer to the coming to be or cessation of a state or determination of some object. For example, the freezing of water is an event because it involves the coming into existence of a new state (solidity) of the water, which is, of course, at the same time also the cessation of its previous state (liquidity). Accordingly, as Kant makes clear in the second edition, every event is an alteration and every alteration an event.

The conception of an objective succession is trickier, since it contains a twofold ambiguity. The first concerns the referent of 'succession', which can refer either to a succession of events (for example, the motion of one body followed immediately upon impact by the motion of another) or to the succession of states constituting a single event (for example, the alteration of the water from a liquid to a solid state). Although Kant is notoriously slippery on this point, in different contexts characterizing both types of succession simply as *A-B*, it seems clear from his formulation of the principle in both editions, as well as from the arguments offered in support of it, that his main concern is with the latter.<sup>27</sup> In other words, the claim is that the concept or, better, schema of causality must be presupposed as a condition of the determination of the succession of states constituting an event, not that one must assume a causal connection between events in order to determine their temporal order.

The second ambiguity concerns the sense of 'objective' and its correlate 'subjective', which suggests two quite distinct contrasts. One is between a real and a merely apparent succession of states (for example, the contrast between real and apparent motion); the other is between a succession of perceptions that is taken (whether correctly or incorrectly) as perceptions of successive states and a succession that is not (for example, the successive perceptions of a house). Depending on which option one chooses, the goal of the Second Analogy is to demonstrate that the schema of causality is required in order to distinguish either between actual events and merely apparent ones or between events and static states of affairs.

The first alternative appears to have some direct textual support. Thus, in concluding his main line of argument, Kant remarks that the relation of cause and effect "is the condition of the objective validity of our empirical judgments with regard to the series of perceptions, thus of their empirical truth, and therefore of experience" (A202 / B247). This reference to the "objective validity" and "empirical truth" of our judgments regarding temporal succession strongly suggests that Kant has in mind the condition of the correctness of such judgments rather than merely the condition of making them at all. Moreover, the previously noted connection between the concept of causality and event types seems to provide indirect support for this reading by indicating the inseparability of the concept of causality from that of a causal law. If one assumes that to apply the schema of causality is to subsume appearances under a causal law, then it is also natural to assume that its function is to distinguish genuine causal sequences from merely apparent ones.<sup>28</sup>

Nevertheless, further consideration suggests that such a reading is both mis-

taken and the source of a good deal of confusion regarding the interpretation of the Second Analogy. As Graham Bird points out, Kant's concern in the Second Analogy is to show that the concept or schema of causality is necessary for taking successive representations as representations of an event, not that it is required for distinguishing between genuine and merely apparent events.<sup>29</sup> Consequently, the question of whether in a particular case one is actually experiencing an event (as opposed to merely imagining one) or experiencing the very event that one takes oneself to be experiencing is an empirical matter, to be decided by the usual empirical means, rather than by appeal to a transcendental principle.

Although a fuller justification of this claim must await an analysis of the actual argument, it is possible at this point to note three considerations that support it. First, as we saw in chapter 7, Kant's claim in §19 that judgment as such possesses objective validity is not to be taken to mean that every judgment is true, only that it has a truth value. Therefore, it seems reasonable to take Kant to be using the notion in the same sense in the present context. Nor is such a reading precluded by the reference to "empirical truth," since Kant claims explicitly that his concern in the Analogies is only with the "formal conditions of empirical truth" (A191 / B236). In the case of the Second Analogy, this is most plausibly taken to refer to the *a priori* framework for judgments regarding the occurrence of an event, as opposed to the "material conditions" required for the determination of the truth of particular judgments. Second, this reading corresponds to the problematic on which Kant himself focuses in distinguishing between the perception of an event (the ship moving downstream) and the successive perceptions of aspects of an enduring object (the house). Third, and most important, even if one is mistaken in one's judgment of an event (as in the case of an illusion or hallucination), it still follows, according to Kant's analysis, that appeal is made to the schema of causality. Thus, the question of its transcendental function must be distinguished from that of the correctness of its empirical application in particular instances.

## B. The Essential Argument

The first problem confronting any interpretation of the Second Analogy is that Kant seems to offer a plethora of proofs. Indeed, commentators have distinguished as many as six, including the one added in the second edition.<sup>30</sup> Moreover, though it is generally recognized that five of these proofs are variations on a single theme, it is sometimes thought that one of them, the so-called argument from the nature of time (A199–201 / B244–46), stands apart from the

rest and involves peculiar difficulties.<sup>31</sup> Nevertheless, these complexities will be ignored here, since the text provides a single line of argument that is contained in both editions and can be expounded and analyzed without appealing to any of the unique and especially problematic features of that particular argument.<sup>32</sup>

The argument begins with a description of the essential features of event perception. Kant remarks: "That something happens, i.e., that something or a state comes to be that previously was not, cannot be empirically perceived except where an appearance precedes that does not contain this state in itself" (A191 / B236–37). In other words, I cannot be aware that "something has happened" unless I can contrast the present state of some object with its preceding state. This much is already clear from the argument of the First Analogy, and so is the consequence that "[e]very apprehension of an occurrence is therefore a perception that follows another one" (A192 / B237). But this is only a necessary and not a sufficient condition of the perception of an event. Since all apprehension is successive, *every* perception follows upon a preceding one. Consequently, the problem is to determine the conditions under which a succession of perceptions can be taken as the perception of a succession of states in the object, that is, of an event.

Kant introduces the example of the perception of a ship sailing downstream in order to clarify the problem. His point is simply that in contrast to the perception of a static object (a house), where I do not regard my successive perceptions as perceptions of a change or succession in the object itself, in the perception of a ship I do view my perceptions in just this way. Kant concludes from this that in the latter case I am constrained to regard the order of my perceptions as determined or irreversible. In other words, if I judge that I am perceiving a change in the position of the ship from point *A* at  $t_1$  to point *B* at  $t_2$ , then I must also think the order of my perceptions as determined, that is, I must think this order as *A-B* rather than *B-A*. One can, of course, imagine a different order of perceptions; but doing so one is imagining a different event, for example, a ship sailing in the opposite direction.<sup>33</sup>

Unfortunately, Kant's manner of characterizing this irreversibility is somewhat misleading. In one place he writes: "The order in the sequence of the perceptions in apprehension is therefore here determined, and the apprehension is bound to it" (A192 / B238). This and similar passages have often led commentators to take Kant to be claiming that in the perception of an event the actual subjective order of perceptions (the order of apprehension) is rendered necessary by the successive order of the states perceived. In other words, the claim

that “apprehension is bound to it” is taken to mean that *its* order is causally determined, and that this is what makes it impossible for the perception to occur in the reverse order.<sup>34</sup>

This would, however, be a strange line of argument for Kant to adopt. First, by treating the schema of causality as, in effect, a rule of apprehension, it ignores the distinction between perception (apprehension) and experience on which the distinction between the Mathematical and Dynamical Principles turns. Second, it cannot be determined by inspection that an order of perceptions is (or is not) irreversible. This is a matter of “interpretation” (albeit not the imaginative variety) and requires subjecting the order to an *a priori* rule. Finally, if irreversibility is understood in this way, then Kant must be taken as claiming that it somehow functions as an “inference ticket,” licensing judgments about the objective temporal order on the basis of a perceived subjective order. But if this is the case, then Kant is hopelessly confused, since it would commit him to the very empirical idealism he adamantly rejects.<sup>35</sup>

How, then, are we to understand this irreversibility? The short answer is that it characterizes the way in which we connect perceptions in thought (the objective unity of apperception), insofar as we represent *through* them an objective succession.<sup>36</sup> In other words, irreversibility does not refer to a given perceptual order, which we can inspect and then infer that it is somehow determined by the object; it refers rather to the conceptual ordering of the understanding through which it determines the thought of an object (in this case objective succession). Prior to the conceptual determination there is no thought of an object at all and, *a fortiori*, no experience.

The task, therefore, is to determine the condition under which we can think an order of perceptions as irreversible and thus as perceptions of an event. Given Kant’s analysis, such a condition can be supplied only by an *a priori* rule. Moreover, since the order in question is temporal, the rule must have the status of a transcendental schema. We need, then, only ask which schema is involved in the thought of such an order, and the answer is clear, namely, that of the pure concept of causality, that is, “the succession of the manifold insofar as it is subject to a rule” (A144 / B183). In fact, this schema is just the form of the thought of an event. Consequently, it is only by subsuming our perceptions under this rule that we can regard them as containing the representation of an event.

The final step is the move from the subsumption of perceptions to the rule to the subsumption of the event itself. Although this is clearly the key step, since it is required to secure the objective reality of the rule, it is also the one most often misunderstood. On a fairly common reading, it is thought that the most

that this line of argument can establish is the necessity of subsuming one's perceptions under the rule, which has no bearing on the rule-governedness of the event itself. Consequently, the only way to preserve the argument is, in phenomenalistic fashion, to reduce judgments about objects to judgments about our *perception* of objects.<sup>37</sup>

It should be clear by now, however, that this cannot be the correct reading. The whole problematic of the Second Analogy is grounded in the assumption of the impossibility of simply identifying the order of perceptions with the order of the successive states of the object perceived. Once again, then, the subsumption of perceptions under a rule cannot be construed as the means for making the perceptions themselves into objects, but rather as the basis for conceiving of a distinct, objective temporal order in and through these perceptions. But in so doing we are necessarily thinking the objective order in accordance with the rule. Indeed, the principle at work here is one that is fundamental to the Analytic as a whole: "The conditions of the **possibility of experience** in general are at the same time conditions of the **possibility of the objects of experience**" (A158 / B197).<sup>38</sup>

This, then, is the central line of argument of the Second Analogy. It is not an argument from the nature of time, though it is concerned with the conditions of the representation of a succession in time. It is rather an argument from the nature of event perception to the conditions of its possibility and can be broken down into five steps: (1) All event perception requires successive perceptions of an object. (2) But this is merely a necessary and not a sufficient condition of event perception. The latter also requires the perception of successive states or determinations of the object, and (since all apprehension is successive) this can never be determined on the basis of the successiveness of the perceptions themselves. (3) In order to consider a succession of perceptions as perceptions of successive states or determinations of an object, it is necessary to regard their order as irreversible. (4) To regard perceptions in this way is just to subsume them under an *a priori* rule, which in this case must be the schema of causality. (5) As a condition of the possibility of the experience of an objective succession, the schema is also a condition of the succession itself (as an object of possible experience).

### C. Evaluation of the Argument and Response to Criticisms

Does the argument, so construed, provide an adequate answer to Hume's challenge to the causal principle? Apart from global doubts about the underlying

transcendental framework, one reason to think that it does not is the apparent weakness of its central claim. If, as has been suggested, the argument shows merely that the schema of causality is necessary for the thought of an event but does not guarantee the veracity of any empirical judgment concerning events, then how can it answer a skeptic regarding causality? At best it seems to shift the skeptic's challenge from the causal principle itself to the judgment of the events that it supposedly conditions.<sup>39</sup>

Such criticism, however, misconstrues Kant's target. It assumes that what Kant tried to do in the Second Analogy (or at least ought to have tried to do) is to answer a Cartesian-type skepticism regarding the "objectivity" of our experience of events. But, as we shall see in the next chapter, such Cartesian-type worries are the concern of the Refutation of Idealism and have no direct bearing on the question before us. For Hume does not doubt that we perceive events, merely that we have grounds (apart from a psychological propensity) for assuming any necessary connection between them. In fact, he must presuppose that we do so and, therefore, distinguish between a mere sequence of perceptions and the perception of an event, in order to offer his well-known account of how we come to form the belief that future sequences of events will resemble past sequences. But since this presupposition is necessary for his own program, Hume must likewise accept anything that is shown to be its necessary condition, which is just what Kant has presumingly shown with the schema of causality.<sup>40</sup>

In response to this it is sometimes insisted that we "just see" an objective succession (that *B* follows *A*) and, therefore, have no need to import any *a priori* principles or transcendental apparatus to account for its possibility.<sup>41</sup> This response derives its plausibility partly from the erroneous assumption that Kant is describing a psychological process whereby one supposedly first inspects one's perceptions and then concludes that they are perceptions of an event, and partly from the previously noted ambiguity in the notion of an objective succession. If it is taken to mean that we can determine that event *B* followed event *A* without having to assume any causal connection between them, this can readily be acknowledged without detriment to the argument. For the latter addresses the more basic question of how we perceive *A* and *B*, taken as discrete events. And to say that these, too, are "just seen" begs the question.

To counter the denial that the Second Analogy is concerned with the conditions of the temporal ordering of successive events rather than the ordering of the stages constituting a single event, the critic might be tempted to appeal to Kant's admission of the frequent simultaneity of cause and effect (A203 / B248).

Kant introduced this issue after the completion of his argument because he apparently saw such simultaneity as the basis for a possible objection, and he addresses it by distinguishing between the *order* and the *lapse* of time. In brief, his claim is that even if, as is usually the case, there is little or no lapse of time between the beginning of the “causality of the cause” and the coming into being of the effect, the former is still prior in the order of time to the latter (A204/B249). But since cause and effect are distinct events, it might be thought that by focusing on their temporal relation, Kant is indicating that the concern of the Second Analogy is, indeed, to secure the ordering of distinct events by subsuming them under the relation of cause and effect. As Guyer has pointed out, however, the temporal order with which Kant is concerned is not that of cause and effect but of the states of the object undergoing alteration in the event.<sup>42</sup>

As Guyer also notes, this of itself is sufficient to dismiss Schopenhauer’s attempted *reductio* of Kant’s argument.<sup>43</sup> Taking Kant to be arguing that the only succession we can regard as objective is that of cause and effect, Schopenhauer quite correctly pointed out that appearances can perfectly well follow *after* one another without following *from* one another. A case in point is the sequence of night and day, which Schopenhauer presumably borrowed from Reid, who used it against Hume.<sup>44</sup> Since Kant (on this interpretation) must deny that there can be any objective yet noncausal succession, Schopenhauer contended that he falls into the error opposite to Hume’s. Just as Hume falsely took “*alles Erfolgen für blossen Folgen*,” so Kant maintains “*dass es kein anderes Folgen gebe, als das Erfolgen*.”<sup>45</sup>

It is clear, however, that this line of objection is based on a complete misrepresentation of Kant’s argument. Since the objective succession with which he is concerned is that of the states constituting an event, the argument has no direct bearing on the objectivity of the succession of distinct events. Consequently, it does not entail that the only succession (of distinct events) we can experience is that of cause and effect. Moreover, returning to Schopenhauer’s example, if we take the succession of night and day to refer to a succession of states of the earth (or a portion thereof), then it is itself an event in Kant’s sense and, therefore, subject to the causal principle.

Finally, we must consider the notorious non-sequitur objection, which likewise assumes that the perception of an event is immediate and non-problematic. This objection was first formulated by Arthur Lovejoy and was later reiterated, apparently without any awareness of Lovejoy’s analysis, by Strawson, who famously termed it “a *non sequitur* of numbing grossness.”<sup>46</sup> For the sake of brevity, we shall here examine only Strawson’s version.



Strawson grants Kant's claim that in the experience of an event the order of perceptions is determined or irreversible and "in this sense necessary," because he takes it to be an attempt to articulate the denial of "order indifference" that is characteristic of event perception. Consequently, Strawson admits that, apart from some possible exceptions, in the perception of a succession *A-B* it is "necessary that the perception of a second state *B* follows and does not precede the perception of the first state *A*." Kant's problem, however, is that he erroneously believes that

to conceive this order of perception as necessary is equivalent to conceiving the transition or change from *A* to *B* as *itself* necessary, as falling, that is to say, under a rule of law of causal determination; it is equivalent to conceiving the event of change or transition as preceded by some condition such that an event of that type invariably and necessarily follows upon a condition of that type.<sup>47</sup>

Here, according to Strawson, lies the non sequitur, which he traces to Kant's illicit and unwitting shift from a conceptual to a causal notion of necessity. It is conceptually necessary that in the perception of the sequence of states *A-B* the observer's perceptions should follow the order: perception of *A*, perception of *B*. Nevertheless, he insists,

the necessity invoked in the conclusion of the argument is not a conceptual necessity at all; it is the causal necessity of the change occurring, given some antecedent state of affairs. It is a very curious contortion indeed whereby a conceptual necessity based on the fact of a change is equated with the causal necessity of that very change.<sup>48</sup>

This whole line of objection is vitiated, however, by the transcendently realistic standpoint it assumes. Ignoring the framework in which the problem is posed (its "transcendental setting"), it treats Kant as if he were an empirical idealist, concerned to ground a conclusion regarding the causal relations of ontologically distinct things and events on a feature of our perceptions (their irreversibility). As Beck points out, Kant could very well respond that, given *this* conception of an object, we could not infer anything about the objective order *A-B* from the order of our perceptions.<sup>49</sup> Moreover, in order to avoid this problem, Strawson must assume that we directly experience events, which, as already noted, begs the question at issue.

In fact, precisely because of his transcendently realistic standpoint, Strawson criticizes Kant for an inference he does not draw, namely, from subjective (in the empirical sense) perceptions to objective (in the transcendental sense) events. Contrary to his assumption, Kant's argument does not move from the

irreversibility of perceptions in a putative instance of event perception (the presumed conceptual necessity) to “the causal necessity of the change.” Rather, as I have already noted several times, the claim is that the schema of causality provides the rule through which we think an order of perceptions as irreversible and thereby take them as perceptions of successive states of an object. In other words, the move is not from irreversibility to causality (that would be a non sequitur), but *to* the assumed irreversibility (in a particular instance) of the sequence of perceptions *by* subsuming this sequence under the schema of causality, *through* which one takes this sequence as the cognition of an event. Once again, the schema is the form of the thought of an objective succession, since it is the rule through which we take the succession of perceptions as yielding the perception of an event.

#### D. An Interpretive Problem:

##### The Schema of Causality or Causal Laws?

The preceding analysis of the Second Analogy may be termed a “weak interpretation,” since it insists that the argument shows (and is intended to show) merely that every event falls under the schema of causality rather than under particular causal laws. As such, it is opposed to a “strong interpretation,” which affirms the latter. Thus, in concluding this discussion, we may find it instructive to reconsider the strong interpretation, which comes in two chief forms.

One of these is the epistemological reading offered by Guyer. According to him, the goal of the Second Analogy is to provide the conditions for confirming particular judgments regarding the occurrence of events and to show that this requires appealing to particular causal laws. As he graphically puts it:

[I]t is only if we are in possession of causal laws which dictate that in the relevant circumstances—that is, not in general, but in the particular circumstances of wind, tide, setting of the sails, and so forth, which are assumed to obtain—the ship could only sail downstream that we actually have sufficient evidence to interpret our representations of it to mean it is sailing downstream.<sup>50</sup>

In addition to his insistence on a tight connection between the Second Analogy and the Refutation of Idealism, Guyer’s account appears to be based on the premise that the objective validity of judgments about events is equivalent to their empirical truth.<sup>51</sup> But, quite apart from these questionable assumptions, there are two main reasons why his reading is highly dubious. First, it is implausible on the face of it, since we are obviously able to recognize many instances of objective succession without being able to subsume that succession

under a causal law. For example, is it really the case that there can be no “objective experience” of water freezing apart from a knowledge of the causal conditions of this change? Surely human beings have experienced (and continue to experience) countless tokens of this event type without knowledge of its conditions, that is, without being able to subsume the event under the appropriate causal law.

The second difficulty is one that Guyer himself recognizes and attempts to counter, namely, that his reading elevates empirical laws, which must be based on experience, into necessary conditions of experience. Guyer’s response turns on a deflation of being a condition of possible experience to being a condition of the *confirmation* of particular judgments regarding the temporal order. Eschewing any metaphysical reading of the Analogies, according to which they function to constitute an objective temporal order, he insists that they are concerned exclusively with determining the epistemological framework for the confirmation of such judgments. Within this purely epistemological framework, he maintains, there is no incoherence or circularity in the notion that causal laws are required to justify claims about objective succession, even though it is only by means of a knowledge of such succession that we acquire knowledge of causal laws. A problem would arise only if the very same sequence is both derived from a particular causal law and employed as evidence for that law.<sup>52</sup>

This response is unsatisfactory because of its myopic view of the interpretive possibilities. The only options Guyer countenances are a strongly metaphysical reading, according to which the human mind somehow constitutes the order of nature by imposing its *a priori* forms upon it (which reflects his understanding of transcendental idealism), or a narrow empiricist reading, according to which the function of the causal principle is to specify the need for particular causal laws in order to confirm or verify judgments concerning the occurrence of events that are initially made without any appeal to these laws. Clearly, however, there is room for at least one other option, namely, to view the Second Analogy as concerned with the form of the thought of an objective succession, the condition for which is supplied by the schema of causality. On this reading, the schema provides the transcendental condition for that initial experience of an event, the confirmation of which is supposedly provided by subsuming the putative event under empirical laws. As has been noted more than once, this is the basic task of the Second Analogy.

A more orthodox version of the strong reading takes the goal of the Second Analogy to be to prove that every event falls under *some* empirical causal law,

the precise nature of which must be learned from experience. Otherwise expressed, the assumption is that Kant there endeavors to reply to Hume by demonstrating a “law of causality” or “principle of induction” that would guarantee particular generalizations or laws.<sup>53</sup>

Advocates of this reading generally recognize that Kant is not committed to the wildly implausible view that in any change of state  $A$ - $B$ , the initial state must seen as the cause or initiating condition of its successor, since this only occurs in relatively few cases.<sup>54</sup> Instead, it is assumed that Kant’s view is that the sequence of states  $A$ - $B$  must be “lawlike” in the sense that, given some initiating condition, the transition from  $A$  to  $B$  is subsumable under a causal law. Schematically expressed, the claim is that for every object of type  $x$  that changes from state  $A$  at  $t_1$  to state  $B$  at  $t_2$ , there must be some “initiating condition”  $C$ , which being given (together with certain unspecified “standing conditions”), states of  $A$ ’s *type* will necessarily be followed by states of  $B$ ’s *type* in all objects of  $x$ ’s *type*. Textual support for this reading is provided by the previously noted connection between the concept of causality and causal law. Since to claim that  $A$  is the cause of  $B$  is to affirm a causal law connecting  $A$ -type events to  $B$ -type events, it might be thought that the conformity of all events to empirical causal laws follows directly from their subjection to the transcendental “law of the connection of cause and effect.”<sup>55</sup>

Against this reading, it must be insisted that the Second Analogy supposedly provides us with a warrant to search for the cause of any event and, therefore, for the causal law under which it may be subsumed. But it does not determine what the cause is or guarantee that we shall be able to discover either it or the relevant causal law. In fact, Kant makes this quite clear in his explanation of the choice of the term *analogy* for the principles falling under that name. As we saw in chapter 8, the analogy is between the two-term relation expressed in the category and its schema, on the one hand, and the relation of a given appearance to some unspecified relatum, on the other. Thus, Kant suggests that the relation provides “a rule for seeking it [the fourth member] in experience, and a mark for discovering it there” (A180 / B222). What turns out to be crucial, however, and serves to distinguish analogies in philosophy from those in mathematics, is that the former provide only a rule or decision procedure for seeking the fourth member, rather than the fourth member itself (A179–80 / B222). Applying this to the causal relation, we see that the Second Analogy enables us to determine *a priori* that for any given event  $y$ , there must be some antecedent event  $x$  from which  $y$  follows in accordance with a rule, but it does not guarantee that it can

be found. Otherwise expressed, it does not ensure that we shall be able to distinguish between merely contingent regularities and genuine causal connections.

By contrast, the Mathematical Principles supposedly relate directly to intuition and ensure that all appearances (qua intuited) are extensive and intensive magnitudes. In the case of intensive magnitudes, Kant illustrates this by noting that we can “determine *a priori*, that is, construct, the degree of sensation of sunlight out of about 200,000 illuminations from the moon” (A178–79 / B221). The point is not that we can know *a priori* that the sensation of sunlight stands in precisely that (or any other) determinate relation to the illumination of the moon. It is rather that we can know *a priori* that it must have some determinate degree and that there is a mathematically constructible ratio between it and that of any other body (in this case the moon). In other words, the principle guarantees the possibility of determining this ratio mathematically. This is precisely what is lacking, however, in the case of the Analogies. Since the latter are concerned with the existence of appearances in time and since, as Kant puts it, “existence cannot be constructed [determined *a priori*]” (A179 / B222), it follows that we can know *a priori* only that an appearance must stand in a necessary relation to some other appearance, but not that we will be able (even in principle) to determine what that other appearance is and the law connecting them.

Returning to the Second Analogy, the obvious question that arises at this point is why, since we supposedly know that every occurrence must fall under a causal law, there remains a substantive issue regarding the determinability of such laws and of particular causes. Granted, their discovery is an empirical matter and may involve great difficulty; but it might still be thought that the Second Analogy is supposed to provide a transcendental guarantee that such laws are there to be found, that is, are discoverable in principle, much as the Mathematical Principles purport to show that any appearance is measurable in principle and the ratio between any two (or more) appearances is determinable on the basis of a common metric.

Although this is an important question, it is not addressed in the Transcendental Analytic. Kant does, however, treat it briefly in the Appendix to the Transcendental Dialectic, in connection with the hypothetical (inductive) use of reason, and more expansively in the Introductions to the third *Critique*, in connection with his account of reflective judgment. The former will be treated in chapter 15, but the latter lies beyond the bounds of this study.<sup>56</sup> Nevertheless, it may be appropriate to close this discussion by citing a passage from the

later work in which Kant characterizes the problem in particularly striking fashion:

For it may certainly be thought that, in spite of all the uniformity of things in nature in accordance with the universal laws, without which the form of an experiential cognition in general would not obtain at all, the specific diversity of the empirical laws of nature together with all their effects could nevertheless be so great that it would be impossible for our understanding to discover in nature an order that we can grasp, to divide its products into genera and species in order to use the principles for the explanation and the understanding of one for the explanation and comprehension of the other as well, and to make an interconnected experience out of material that for us is so confused (strictly speaking, only infinitely manifold and not fitted for our power of comprehension). [KU 5: 185; 72]<sup>57</sup>

As this passage indicates, the problem is that the transcendental analysis of the first *Critique*, which guarantees the uniformity (lawfulness) of nature at the most general level, does not also guarantee its uniformity at the empirical level. In spite of everything that the Analogies have shown, it remains possible that nature is so complex that the human understanding might never be able to find its way about in it. In other words, Kant introduces another specter, which is distinct from that underlying the Transcendental Deduction and which cannot exorcized by its means.

#### IV. THE THIRD ANALOGY

The Third Analogy is identified in the first edition as the “Principle of community” (A211) and in the second as the “Principle of coexistence [*Zugleichsein*], according to the law of interaction, or community” (B256). Like its predecessors, it is formulated differently in the first and second editions. In the first it states: “All substances, insofar as they coexist, stand in thoroughgoing community (i.e., interaction with one another)” (A211); in the second it reads: “All substances, insofar as they can be perceived in space as coexisting, are in thoroughgoing interaction” (B256). Even though the second edition formulation differs from the first both in its more explicitly epistemological orientation (referring to the perception of coexisting substances) and in its apparently narrower scope (limited to substances in space), there does not appear to be any substantive difference between them. Already in the first edition the argument is concerned with the conditions of the experience of coexistence, and in both editions it is assumed that the only substances whose temporal relations can be determined

are those existing in space. Thus, the basic claim is that thoroughgoing interaction is the condition under which distinct substances can be experienced as co-existing in a common spatiotemporal world.

Since this is a bold claim, whose roots extend to some of Kant's earliest physical and metaphysical speculations, its relative neglect in the literature may seem surprising. Nevertheless, it is understandable (if not justified), given the extraordinary attention paid to the Second Analogy as the locus of Kant's "answer to Hume," the obscurity of its argument, the already noted problematic nature of the connection between the category of community and the disjunctive form of judgment, and, most of all, the appearance of redundancy, stemming from the fact that, like the Second Analogy, it too deals with causality, albeit of the reciprocal variety.

Although the obscurity of the argument and the non-obvious nature of the connection between category and judgment form cannot be gainsaid, the Third Analogy is not made redundant by the Second. On the contrary, it will be argued that it has a distinct and equally essential function as a condition of experience. The discussion is divided into four parts and largely mirrors that of the Second Analogy. The first attempts to clarify some of the terminological issues that account for a good deal of the obscurity. The second analyzes the argument, focusing on the tighter formulation in the second edition. In the light of this analysis, the third explores and responds to the redundancy objection. Finally, the fourth attempts to clarify further Kant's central claims, particularly as they bear on the conception of community.

### A. Some Preliminaries

In the order of exposition, the first of numerous terminological-conceptual problems posed by the Third Analogy is the proper rendering of *Zugleichsein*, which can be translated either as 'coexistence' or 'simultaneity'.<sup>58</sup> Kemp Smith generally translates it in the former way, and Guyer and Wood in the latter. Even though Kant occasionally also uses the latinate *Koexistenz*, thereby suggesting a contrast with *Zugleichsein*, the former better accords with the main thrust of the argument. The basic point is that simultaneity is a relation between events or states of affairs at a particular moment of time, whereas coexistence refers to a diachronic relation between substances that exist throughout a common period of time.<sup>59</sup> As we shall see, in the Third Analogy Kant is concerned primarily (though not exclusively) with the latter.

A frequently noted problem is that since the First Analogy has purportedly already shown that all substance is permanent, the coexistence of substances

(throughout all time) can be inferred without any appeal to a distinct principle. As Paton points out, however, Kant is presumably concerned with the conditions of the empirical knowledge of the coexistence of substances qua in certain states or in respect to certain accidents.<sup>60</sup> In other words, the concern is with the coexistence of determinate things, which, according to the First Analogy, are themselves accidents of substance or, equivalently, ways in which the substances of which they are determinate forms exist. Consequently, they need not be viewed as persisting throughout all time.

The second term requiring clarification is 'interaction' (*Wechselwirkung*). Kant's language in one passage suggests that he construes it simply as a double or reciprocal causality (A212 / B259). For the most part, however, he takes it as equivalent to the relation of 'reciprocal influence' (*wechselseitiger Einfluss*). Since 'influence' is a technical term taken from Baumgarten and refers to the action of one substance upon another, a reciprocal influence would be a reciprocal action between two (or more) substances.<sup>61</sup> In a broad sense, such action is undeniably a kind of causality, which explains Kant's use of causal terminology. As Eric Watkins points out, however, it differs from the event-event causality treated in the Second Analogy, since it consists in a direct relation between substances, which may be thought of as continuous and, therefore, need not involve a change of state, that is, an event.<sup>62</sup> The paradigmatic example of this relation is the continuous exertion of force (attractive and repulsive) of interacting substances.

This understanding of interaction also provides the key to modeling it. If one conceives of reciprocal influence in terms of the model suggested by the Second Analogy, one would view the influence as holding between successive states of the substances involved. Using Kant's illustration of the earth and moon ( $E$  and  $M$ ), on this model  $E_1$  is a cause of  $M_2$  and  $M_1$  a cause of  $E_2$ , and so forth. But in addition to whatever conceptual problems this model may involve, it cannot account for coexistence. For, as Watkins has also pointed out, it requires only that  $M_2$  occurs after  $E_1$  and  $E_2$  after  $M_1$ , and so on; thus it does not show either that  $M_1$  coexists (or is simultaneous) with  $E_1$  or that  $E_2$  coexists (or is simultaneous) with  $M_2$ .<sup>63</sup>

This problem does not arise, however, if we assume that the influence obtains between simultaneous states of the substances:  $E_1$  influences  $M_1$ ,  $E_2$  influences  $M_2$ , and vice versa. Moreover, this model does not conflict with the Second Analogy. The latter is concerned solely with the conditions of the experience of an event, and such influence (again to be thought of in terms of an ex-



ertion of force) does not constitute an event. Accordingly, even though Kant himself does not discuss the issue, it seems reasonable to assume that he had this model in mind.<sup>64</sup>

The final term requiring preliminary discussion is 'community' (*Gemeinschaft*), which Kant presents as an alternative to 'interaction' as a name for the category. Kant remarks that the term is ambiguous and can mean either *communio* or *commercium*, but that here he uses it in the latter sense as referring to a "dynamical community without which even the local community (*communio spatii*) could never be empirically cognized" (A213 / B260). As the latter suggests, the reference to community serves to underscore the unique role of the Third Analogy as a condition of the relation of substances in space. Although Kant's main concern is with coexistence, the fact that the coexistence of substances can be experienced only in space entails that interaction or dynamical community is thereby also the condition of the possibility of the experience of substances as existing together in a single space.<sup>65</sup> In fact, in a letter to J. G. Schulz from 1784, Kant goes so far as to claim that interaction is a condition of outer experience (Br 10: 367; 215).

Nevertheless, it might be wondered why Kant should treat 'interaction' and 'community' as interchangeable, since they clearly are not synonyms. Moreover, the situation is made more puzzling by the fact that Kant also connects the concept of community with the concept of a whole (*Ganze*).<sup>66</sup> For it certainly does not appear to be the case that substances must be thought to constitute a whole in order to be experienced as interacting and/or coexisting. To anticipate, the key to the answer is that the dynamical community is composed of *all* the substances that coexist throughout a stretch of time and only the interaction of the complete set of these substances is sufficient to determine their coexistence. In other words, the coexistence of *E* and *M* is a function not only of their direct interaction but of their interaction with all of the coexisting substances, which collectively constitute a community. Unfortunately, however, Kant obscures this point by his tendency to depict interaction as a relation between two substances.

## B. The Second-Edition Argument

As is the case with the first two Analogies, the argument added in the second edition is contained in a paragraph placed at the beginning of the text (B256–58). The argument can be broken down into seven steps, which, following the usual procedure, will be quoted and discussed in turn.

[Step 1.] Things are **coexistent** if in empirical intuition the perception of one can follow the perception of the other **reciprocally**. . . . Thus I can direct my perception first to the moon and subsequently to the earth, or, conversely, first to the earth and subsequently to the moon, and on this account, since the perceptions of these objects can follow each other reciprocally, I say that they coexist.

Kant here appears to be offering an empirical criterion for the determination of coexistence. According to this criterion things coexist just in case the perceptions of these objects can follow each other reciprocally. If, as previously suggested, Kant is here referring to a diachronic process in which objects are experienced as coexisting through a stretch of time, the situation can be represented schematically as either  $E_1, M_2, E_3, M_4$ , and so forth, or  $M_1, E_2, M_3, E_4$ , and so forth.

The actual situation, however, turns out to be somewhat more complicated. Since the possibility of a reciprocal ordering of our perceptions can no more be read directly off these perceptions than can their irreversibility, there is a problem about how this criterion can be applied empirically. The epistemic situation precisely parallels the one described in the Second Analogy, which means that we cannot use this reciprocal orderability as a premise from which to infer the coexistence of what is perceived.

[Step 2.] Now coexistence is the existence of the manifold at the same time [*in derselben Zeit*].

This is the second edition's definition of 'coexistence', and it closely parallels that of the first (A211). Since this has already been discussed, we need merely point out that the definition (in both editions) refers to the existence of things (or the manifold) *in* the same time, which may be taken to mean either in the same instant or the same stretch of time. In the former case we have simultaneity, and in the latter coexistence. Otherwise all appearances would be coexistent, since they all exist in the same time.<sup>67</sup>

[Step 3.] But one cannot perceive time itself and thereby derive from the fact that things are positioned at the same time that their perceptions can follow each other reciprocally.

This is the ubiquitous premise regarding the unperceivability of time, which plays a central role in all of the Analogies (albeit a somewhat muted one in the Second). What is noteworthy here is that Kant concludes from it that we are unable to determine that our perceptions are reciprocally orderable. Although this is certainly true, one might expect him to have concluded that we cannot

determine coexistence by referring our perceptions to "time itself." Indeed, it is only because time is unperceivable that we have to consider the reciprocal orderability of our perceptions in order to determine coexistence in the first place. But in spite of this awkwardness, it seems reasonably clear that Kant is here attempting to set the stage for the subsequent steps in the argument by highlighting the difficulty of determining such orderability.

[Step 4.] The synthesis of the imagination in apprehension would therefore only present each of these perceptions as one that is present in the subject when the other is not, and conversely, but not that the objects coexist, i.e., that if the one is then the other also is in the same time, and that this is necessary in order for the perceptions to be able to succeed each other reciprocally.

As before, the crucial point is the inability of mere apprehension to determine an objective temporal order (whether of coexistence or of succession). But rather than appealing to the successive nature of apprehension (as he did in the Second Analogy), Kant now remarks that what apprehension discloses is merely that when one of two putatively coexisting things is being perceived the other is not, and vice versa. In other words, at any given time I can apprehend either one or the other but not both. Thus, the problem is that I cannot infer from this that the objects coexist, since it is possible that the latter did not exist when I apprehended the former, and vice versa.

Of course, if I knew that I could apprehend them in either order, then (by step 1) I would also know that the objects coexist. But this is precisely what mere apprehension can never show, since it amounts to the counterfactual claim that at  $t_1$  I *could have* apprehended  $E_1$ , rather than  $M_1$  (or vice versa), even though I actually apprehended the latter.<sup>68</sup> As was the case in the first two Analogies, such a determination requires subjecting my apprehension to a rule that cannot be derived from the apprehension itself.

[Step 5.] Consequently, a concept of the understanding of the reciprocal sequence of the determinations of these things coexisting externally to each other is required in order to say that the reciprocal sequence of perceptions is grounded in the object, and thereby to represent the coexistence as objective.

Here Kant makes explicit what was implicit in the preceding step and characterizes the sought-for rule as a concept of the understanding, more specifically, a concept of the reciprocal sequence of the determinations of the things perceived. The next and crucial tasks are to identify this rule and establish its objective reality. But before turning to these matters there are two points to be

noted, which are essential to relating the argument of the Third Analogy to Kant's theory of logical functions and categories. First, the fact that the rule in question involves reciprocal orderability enables us to see more clearly why it serves for the *coordination* of our representations in a judgment rather than, as with the causal principle, their *subordination*. Second, since only one of the possible sequences can actually be perceived, the judgment of coexistence must take the form of an exclusive disjunction. Again confining ourselves to the diachronic experience of the successive determinations of two coexisting things (the earth and the moon), the rule for thinking coexistence indicates that I can apprehend *either* the order  $M_1, E_2, M_3$ , and so forth, *or*  $E_1, M_2, E_3$ , and so forth, but *not both*.<sup>69</sup> Adding other coexisting objects increases the number of disjuncts, but does not affect the basic point, namely, that the rule for judging coexistence necessarily takes a disjunctive form and requires our regarding the coexisting things as constituting a community.

[Step 6.] Now, however, the relation of substances in which one contains determinations the ground of which is contained in the other is the relation of influence, and, if the latter reciprocally contains the ground of the determinations of the former, it is the relation of community or interaction.

Here Kant specifies the sought for rule or concept of the understanding as that of community or interaction. Combined with the preceding step, this entails that the conception of things as united by this relation is a necessary condition of the representation of them as coexisting externally to one another in space. Presumably, this is the point Kant had in mind when he claimed that interaction is a condition of outer experience.

Once again, the key to the argument lies in the unperceivability of time. Strictly speaking, what makes my perceptions reciprocally orderable is that they are perceptions of coexisting things. If time could be perceived, I could simply notice the coexistence of the objects perceived, from which I could then infer the reciprocal orderability of my perceptions; though in that case the inference would be superfluous. But since time cannot be perceived and the coexistence of objects is not determinable simply by referring them to it, in order to determine coexistence I must be able to interpret a sequence of perceptions as perceptions of successive states of coexisting things, that is, as reciprocally orderable. This is only possible, however, if I regard the objects perceived as reciprocally influencing each other.

The latter point is the nerve of the argument and the target of much of the criticism directed against it. The basic problem is that it seems counter-intu-

itive to suggest that in order to judge two or more material objects as coexisting one must also consider them as interacting. Appealing to the previous analysis of the Second Analogy, it might be claimed that such a conclusion is analogous to the mistaken assumption that in order to experience the succession  $A-B$  as objective it is necessary to consider  $A$  and  $B$  as related as cause and effect. Since this is manifestly false, it might seem that it is likewise false that we need conceive  $A$  and  $B$  as interacting in order to experience them as coexisting.<sup>70</sup>

This comparison with the Second Analogy is highly misleading, however, since it ignores the nature of the objective succession with which the latter is concerned. As we have seen, to think that Kant's argument in the Second Analogy commits him to the absurd consequence that all objective succession is that of cause and effect is a mistake because the succession with which it is concerned is that of the states of a substance constituting an event. In the case of the Third Analogy, however, the relation of coexistence is between distinct substances and their states; so this negative lesson from the Second Analogy is not applicable here.

Moreover, as we shall see in more detail later, Kant distinguishes in the course of the Third Analogy between an immediate and a mediate community. Consequently, he does not in fact claim that it is necessary to conceive two substances as directly interacting in order to experience them as coexisting. All that is required is that each be conceived as standing in such a relation with some co-existing substances, which, perhaps by a complex chain of mediations, stand in this relation to the other.

Finally, it must be kept in mind that the reciprocal orderability of perceptions, through which coexistence is represented, is taken as grounded in the object (otherwise it would not yield objective coexistence), which means that each of the two possible orders must likewise be viewed as so grounded. Thus, if the order actually perceived were  $E_1, M_2, E_3, M_4$  and so forth, I must take both it and its reciprocal as equally objective, that is, as grounded in the relation between  $E$  and  $M$ . But interaction or reciprocal influence is the only relation capable of licensing the thought of this reciprocal orderability. Absent the perceivability of time, which would presumably allow for a direct inspection of coexistence, thereby bypassing any need for a rule, there is simply nothing else capable of doing the job.

[Step 7.] Thus the coexistence of substances in space cannot be cognized in experience otherwise than under the presupposition of an interaction among them; [8] this is therefore also the condition of the possibility of the things themselves as objects of experience.

Although this sentence contains two claims, we may regard it as a single step. First, it makes explicit the previously noted point regarding the epistemic function of the concept or, better, schema of interaction, and then it affirms its objective reality. The latter is the goal at which the argument has been directed all along, but since precisely the same considerations apply here as in the Second Analogy, nothing more need be said about it.

### C. The Redundancy Problem

Assuming the basic framework of Kant's treatment of the Analogies, which entails the necessity of appealing to *some* rule in order to determine coexistence, the predictable rejoinder is to challenge the assumption that interaction is, in fact, the required rule. Specifically, the critic will claim that the task can be performed by the unidirectional principle of causality, which, in turn, leads to the charge of redundancy. Although the locus classicus of this line of objection is Schopenhauer,<sup>71</sup> here we shall focus on two contemporary formulations, those of Melnick and Guyer.

(1) Melnick: Melnick does not deny the distinction between causality and interaction or, as he puts it, between causal laws and laws of interaction, but he does deny that it correlates with a distinction between laws that determine the time order of appearances as successive and those that determine it as simultaneous or coexistent. Thus, he maintains that Kant's attempt to separate his basic argument for the role of causality (and causal laws) into two distinct Analogies is "artificial and forced."<sup>72</sup>

Melnick bases his case on a putative counterexample designed to show that determining succession cannot be correlated specifically with causation and determining simultaneity or coexistence with interaction. As he initially describes the situation,

Suppose we have two billiard balls, *b* and *c*, at rest, and they are hit by a moving billiard ball *a*, whereupon *b* and *c* move off in different directions [diagram inserted]. We thus have two series of successive states: 1) *b* at place *p*<sub>1</sub>, *b* at *p*<sub>2</sub> . . . , *b* at *p*<sub>*n*</sub>; 2) *c* at *p*<sub>1</sub>', *c* at *p*<sub>2</sub>', . . . , *c* at *p*<sub>*n*</sub>'. Again, we have a dynamical law *L*<sub>1</sub> that describes the motion of *b* and *c* as a function of various factors—the elasticity of the billiard balls, the direction of the acting force (the angle at which *a* hit *b*, and the angle at which *a* hit *c*), the magnitude of the acting force (the momentum of *a*), the coefficient of friction of *b* and *c* with respect to the billiard table, and so on.<sup>73</sup>

Given this scenario, Melnick thinks it easy to show that causality can of itself determine coexistence without any appeal to interaction. Reduced to the sim-

plest terms, this is because, given the collision of *a* with *b* and *c* (the initiating condition) together with  $L_1$ , it is perfectly possible to plot the relative positions of *b* and *c* without assuming their effect either on *a* or on each other. Melnick acknowledges a gravitational attraction between *b* and *c*, but he insists that this may be safely disregarded. Consequently, he believes himself entitled to conclude that “the position of *b* at any time is not *directly* [my emphasis] a function of the position of *c* or vice versa; i.e., *b* and *c* are not *in any important sense* [my emphasis] in interaction.”<sup>74</sup> Or, as he also puts it, “We can use the same law ( $L_1$ ) under the same conditions (the collision) to determine both succession and simultaneity of states.”<sup>75</sup>

There are two basic points to be made about Melnick’s analysis. First, he apparently understands interaction to require a direct reciprocal influence or, as Kant puts it, an “immediate community.” Thus, it is questionable whether his analysis really addresses the Kantian position. Second, Melnick is correct to suggest that the place of *b* at a certain point in time (say *n* seconds after its impact with *a*) can be determined independently of its relation to *c* and that of *c* independently of its relation to *b*, since each of these places, *taken individually*, can be determined by subsuming the relevant data under  $L_1$ . His mistake is to assume that this is sufficient to determine their coexistence. The latter requires not merely being able to predict where each will be at a particular point of time but also being able to identify the time at which *b* is at *pt* with the time at which *c* is at *pt*’. But this cannot be done apart from a determination of the temporal relation of *b* and *c* at their respective locations, which, as we have seen, presupposes their reciprocal influence. Consequently, Melnick begs the question at issue by simply assuming coexistence rather than explaining how its representation is possible or, equivalently, what entitles us to take our successive representations as reciprocally orderable.<sup>76</sup>

(2) Guyer: Even though Guyer emphasizes the systematic significance of the Third Analogy, his view of the shortcomings of its argument is much like Melnick’s. According to his reading, the basic argument proceeds as follows: (1) Although the reversibility (reciprocal orderability) of a sequence of perceptions may be the sign that they are perceptions of coexisting states of affairs, we are no more directly given this reversibility than (in the case of the Second Analogy) we were given their irreversibility. (2) Moreover, even if we knew that our perceptions were reversible, this would still not be a *sufficient* condition for judging that the *objects* represented by them coexisted. Presumably, this is because it would not preclude the possibility that the objects themselves were succeeding

each other in the same order as the representations.<sup>77</sup> (3) Consequently, he continues:

If both A and B are to coexist through the whole period of time  $t_1$  to  $t_2$  in which I have, for instance, a succession of representations of the type  $A_r$ -then- $B_r$ , and if *a fortiori* I am to know that instead of that sequence I could then have had representations of the type  $B_r$ -then- $A_r$ , then I must know that at the time  $t_1$  when I was perceiving A . . . I *could have* been perceiving B . . . instead. But my only basis for believing this must be the belief that the state of B at  $t_1$  "determines the position" of A as also existing at  $t_1$ . . . . [For] only a relationship between A and B such that the state of A at  $t_1$  is necessarily connected with the state of B at  $t_1$  will provide me with the evidence necessary to judge that, although it was A that I was perceiving at  $t_1$ , B also existed at  $t_1$  and *a fortiori* could have been perceived by me then. . . . The *ground* for such an inference, in turn, would have to be that the state of A at that time either actually depends on the simultaneous existence of B or else that it must produce it; . . . the state of the one is a necessary condition of the state of the other, thus that the state of one object at one moment must be either necessary cause or necessary effect of the state of the other at that moment.<sup>78</sup>

Guyer's critique is a direct consequence of this reading of the argument. Since, on his view, all that is required to justify an inference to the coexistence or simultaneity of states of affairs is a relation of causal dependence (where one of the states of affairs is taken *either* as cause *or* effect of the other), it follows that a distinct relation of interaction or reciprocal causality is superfluous. Thus, he concludes:

What Kant's argument requires is simply that the existence of one object in one state at one moment necessitates the existence of another object in another state at that same moment, so that it may be inferred that both objects actually exist at the same moment and *a fortiori* that either could be perceived then. But this logical relation of necessitation will be satisfied if one object *either* depends upon *or* produces the state of the other at that moment; it is therefore not obvious that each state must be *both* cause and effect of the other.<sup>79</sup>

There are four points to make in response to this diagnosis of Kant's argument. First, as Guyer himself notes, his analysis turns crucially on the simultaneity of cause and effect. Thus, he makes much of the fact that Kant allows for the possibility of the exercise of the "causality of a cause" being simultaneous with the coming-into-being of its effect. In this case, however, the cause must *necessarily* be simultaneous with its effect, which is certainly a claim that cannot be made about all causes and, therefore, cannot be taken as a consequence of the analysis of the causal principle in the Second Analogy.<sup>80</sup>



Second, Guyer, like Melnick, illicitly helps himself to the notion of simultaneity. Since not *every* cause is simultaneous with its effect, one cannot simply appeal to the relation of causal dependence to determine that two states of affairs are simultaneous. Rather, one must appeal to a relation in which both cause and effect “exist at the same moment,” so that one could perceive either. But since determining the latter is just what requires explanation, Guyer’s attempt to show that the relation of unidirectional causality is sufficient to determine coexistence fails on the face of it.<sup>81</sup>

Third, Guyer’s account completely ignores Kant’s distinction between the relation of subordination thought through the category of causality and the relation of coordination thought through the category of community. Admittedly, this neglect is of no concern to Guyer, since he explicitly dismisses any putative connection between the Analogies and the issues concerning the Metaphysical Deduction. Nevertheless, this dismissal comes at a price, since it leads to the neglect of the connection between the category and the thought of the reciprocal orderability of one’s perceptions, which, as we have seen, amounts to the form of the thought of coexistence.

Finally, by appealing to the relation of causal necessitation as both necessary and sufficient to ground a judgment of coexistence, Guyer commits Kant to the implausibly strong requirement that in order to establish coexistence or simultaneity, it is necessary to subsume the putatively simultaneous items under a causal law, “which dictates that they *must* exist simultaneously.”<sup>82</sup> Given what we have already seen about Guyer’s view on the Second Analogy, this should come as no surprise. Once again, however, it should give one reason to question the adequacy of his reading, particularly given Kant’s distinction between immediate and mediate community, which Guyer, like Melnick, completely ignores.

#### **D. Further Considerations Regarding the Nature of Community and the Strength of Kant’s Claim**

The above considerations point to the need to gain a clearer understanding of the distinction between an immediate and mediate community, which, in turn, leads us back to the concept of community itself. As we have seen, Kant construes this concept in its dynamical sense (as *commercium*), which involves the thought of a whole, each of whose members influence and are influenced by each of the others at every moment throughout an indefinite stretch of time. This concept of community also indicates the profoundly Leibnizian picture of

nature operative in the Third Analogy, one that differs sharply from the more familiar Humean picture of discrete events at work in the Second. In spite of the replacement of a universal and pre-established harmony by interaction, Kant retains the core Leibnizian conception of a functional interdependence linking all parts of corporeal nature.

It is against this backdrop that we must consider the distinction between an immediate and a mediate community to which Kant appeals at various points in the Third Analogy. As a starting point, it is useful to consider the distinction drawn by Eric Watkins between “strong” and “weak” interpretations of Kant’s conception of interaction, which parallels the earlier distinction between strong and weak interpretations of the Second Analogy.<sup>83</sup> According to Watkins, the strong interpretation holds that every substance in the universe directly interacts to some degree with every other, whereas the weak interpretation maintains that the interaction may also be indirect. As Watkins duly notes, the former is strongly suggested by the Newtonian principle of universal gravitation to which Kant appeals in the *Metaphysical Foundations of Natural Science*; but this does not, of itself, preclude the latter or weak interpretation, since it allows for both direct and indirect interaction.

Moreover, the weak interpretation both accords better with the text of the Third Analogy and makes better philosophical sense of its argument. It better accords with the text because, as already noted, Kant himself explicitly distinguishes between an immediate and a mediate community (a distinction that remains completely mysterious on the strong interpretation). It makes better sense of the argument because, as the criticisms of Melnick and Guyer indicate, direct interaction (particularly when interpreted as they do as amounting to reciprocal causal determination) is an implausibly strong necessary condition of the determination of coexistence.

In addition, Kant’s own cryptic discussion of the topic strongly indicates that he viewed the establishment of a merely mediate community as sufficient for the determination of coexistence. Thus, by way of illustrating the doctrine of thoroughgoing interaction he writes:

From our experiences it is easy to notice that only continuous influence in all places in space can lead our sense from one object to another, that the light that plays between our eyes and the heavenly bodies effects a mediate community between us and the latter and thereby proves the coexistence [*Zugleichsein*] of the latter, and that we cannot empirically alter any place (perceive this alteration) without matter everywhere making the perception of our position possible; and only by means of its reciprocal influence can it establish their simultaneity [*Zugleichsein*] and thereby the

coexistence [*Koexistenz*] of even the most distant objects (though only mediately).  
[A213 / B260]

Although this dense passage lumps at least four distinct claims together in a single sentence and leaves unexplained the reciprocal nature of the influence in question, it clearly suggests that a merely mediate community, such as that which is effected by the intervening rays of light (not gravitational force), is sufficient to establish the coexistence of the observer with remote objects in space. To be sure, Kant is concerned here with the determination of an observer's coexistence with such objects, which is perhaps why the influence of the observer on the observed is ignored. Nevertheless, there is no reason to assume that he did not think that the same conception of a mediate community or interaction applies, *mutatis mutandis*, to the coexistence of observed objects with each other as well.

As Watkins also points out, however, this requires that the relation of community and, therefore, interaction, be viewed as transitive.<sup>84</sup> In other words, if *A* stands in a dynamical community or relation of reciprocal influence with *B* and *B* with *C*, then *A* and *C* are similarly related, even though (apart from gravitational attraction) there may be no direct interaction between them. But this is possible if, as suggested earlier, we take the reciprocal influence (whether mediate or immediate) to hold between simultaneous rather than successive states of the coexisting substances.<sup>85</sup> Admittedly, Kant nowhere states explicitly that he takes the relation to be transitive; but we have seen that he must do so if he is to distinguish between an immediate and a mediate community. More important, he must do so if, as the argument contends, this relation is to serve as the condition of the representation of the transitive temporal relations of simultaneity and coexistence.

So construed, the claim of the Third Analogy is "weak" relative to that of the Second, since its conception of influence is far weaker than the conception of cause (as sufficient condition) operative in the latter. In other respects, however, its claim is weak in precisely the same sense. In both cases it is matter of subsuming successive representations under the schema of a category (here "the coexistence of the determinations of the one [of two putatively coexisting substances] with those of the other in accordance with a general rule" [A144 / B183–84]), rather than under a determinate law. Presumably, prior to Newton people were perfectly capable of forming judgments regarding the coexistence of objects successively perceived and, if Kant's argument establishes anything, it is that such judgments must involve subsuming perceptions under a rule that

warrants the thought of their reciprocal orderability. Moreover, as was the case with the Second Analogy, the regulative nature of the principle indicates that we are permitted to search for laws governing interaction, but it neither specifies what they are nor guarantees that we shall find them. Accordingly, since we have experience of objective coexistence (as well as succession) independently of such laws, the latter cannot be viewed as necessary conditions of the possibility of experience. Finally, the Third Analogy is not to be taken as providing an argument for Newton's third law of motion (the equality of action and reaction in all communication of motion), which Kant characterizes as the third mechanical law and attempts to establish in the *Metaphysical Foundations of Natural Science* (MAN 4: 544–47; 252–54). Like the law of the conservation of the quantity of matter, the latter is an *application* of the transcendental principle rather than an explication of it.<sup>86</sup>

## Chapter 10 Inner Sense and the Refutation of Idealism

This chapter is concerned with two closely related topics, which Kant brought together in the second edition of the *Critique*: the theory of inner sense and the refutation of idealism. To be sure, these topics are not treated there for the first time. The conception of an inner sense and of time as its form, as well as the associated doctrine of the phenomenality of self-knowledge, are already found in the first edition. And we have already seen that the Fourth Paralogism of the first edition contains a refutation of idealism. Nevertheless, it is only in the second edition that Kant thematizes the problem of inner sense, links it to the conception of self-affection, and attempts to explain how the latter grounds the merely phenomenal character of self-knowledge. It is likewise in the second edition that Kant first attempts to base his refutation of problematic idealism on the thesis that cognition of outer objects is a condition of the empirical cognition of one's own existence as determined in time. Accordingly, the present chapter falls naturally into two parts: the first treating inner sense and its connection with self-knowledge, and the second the refutation of problematic idealism based upon it.

## I. INNER SENSE AND SELF-KNOWLEDGE

Few would quarrel with Paton's lament that "Kant's doctrine of self-knowledge is the most obscure and difficult part of his philosophy."<sup>1</sup> Moreover, the reasons for this are not hard to find. Ultimately, they stem from the fact that Kant's account of self-knowledge is rooted in his theory of inner sense, according to which we can know ourselves only as we appear to ourselves. Not only is this theory of the sensory nature of self-knowledge inherently paradoxical, indeed, it is presented as such by Kant, but his scattered discussions of the topic are exceedingly fragmentary. These are based on the conception of time as the form of inner sense, and their aim is to show that self-knowledge is subject to the same transcendental conditions as the knowledge of objects of outer sense. The present discussion will focus on this general thesis and is divided into three parts: (1) an analysis of the claim that time is the form of inner sense; (2) a determination of the nature of the object of inner sense and inner experience; and (3) an examination of Kant's argument for the phenomenality of this object.

### A. Time as the Form of Inner Sense

In the Aesthetic, Kant draws three conclusions from the metaphysical and transcendental expositions of the concept of time. The first of these parallels the first conclusion regarding space, namely, "Time is not something that would subsist for itself or attach to things as an objective determination, and thus remain if one abstracted from all subjective conditions of the intuition of them" (A32 / B49). The remaining two, however, are designed to underscore the differences between the status and functions of space and time. The second states, "Time is nothing other than the form of inner sense, i.e., of the intuition of our self and our inner state." In support of this restriction of time to inner sense, Kant remarks that "time cannot be a determination of outer appearances; it belongs neither to a shape or a position, etc., but on the contrary determines the relations of representations in our inner state" (A33 / B49–50). The third conclusion states, "Time is the *a priori* formal condition of all appearances in general" (as opposed to space, which is merely the pure form of all outer intuitions) (A34 / B50).

At first glance, the second and third conclusions appear to conflict with one another, since the former seems to limit the scope of time to inner appearances, while the latter asserts that it is a condition of *all* appearances. Kant attempts to show that this conflict is merely apparent, however, by stating that time is "the immediate condition of the inner intuition (of our souls), and thereby also the

mediate condition of outer appearances.” It is the latter because all appearances, qua representations or modifications of the mind, are given in inner sense. Moreover, since time is the form of inner sense, this enables Kant to conclude that “all appearances in general, i.e., all objects of the senses, are in time, and necessarily stand in relations of time” (A34 / B51).

The full import of these remarks first becomes apparent in the Analogies. As we saw in the last chapter, the objective temporal order is thought rather than intuited, which is a direct consequence of Kant’s denial that time is a determination of outer appearances, that is, a quality of such appearances analogous to their size, shape, spatial position, and sensory properties.<sup>2</sup> Inner sense was there viewed as providing a merely subjective order of the succession of representations in empirical consciousness and the need for *a priori* principles of objective time-determination was seen to stem from the fact that an objective order can be determined neither by inspecting this subjective order (since it is always successive) nor by directly appealing to an objective temporal order (since time cannot be perceived). Now, by contrast, the focus shifts to inner sense as a sensory form of self-awareness, through which the mind intuitively itself and its states. Inner sense, so construed, is contrasted not with outer sense but with apperception (or at least “pure” or transcendental apperception). In the *Anthropology* Kant characterizes the latter as a “consciousness of what we are *doing*” and the former as a “consciousness of what we *undergo* insofar as we are affected by the play of our own thoughts (Anthro 7: 161; 39). Otherwise expressed, apperception is an “intellectual consciousness” of the act of thinking (a consciousness of spontaneity), whereas inner sense is a sensory consciousness of the contents of thought.

Consequently, what the mind is aware of through inner sense or, equivalently, introspection, are just its own representations, all of which pertain to outer sense. In other words, inner sense has no manifold of its own, which is why Kant insists that the representation of time requires an appeal to space. Thus, in the second conclusion, immediately after denying that time can be a determination of outer appearances, since it “belongs neither to shape or position,” Kant remarks: “And just because this inner intuition yields no shape, we also attempt to remedy this lack through analogies” (A33 / B50). Not surprisingly, these analogies turn out to be spatial. Specifically, we are constrained to represent time in terms of a line progressing to infinity, with the proviso that the parts of a line coexist, while those of time are successive. When we first encountered this thesis in chapter 7 it was in connection with an analysis of the transcendental function of the imagination. As we saw there, the representa-

tion of time through a line involves an imaginative “interpretation” of this line, which by attending to the act through which it is generated rather than the product of this act, the line is taken as something other than it immediately presents itself as being. But the inherently spatial nature of the contents of inner sense and of the representation of time, the form of inner sense, also have implications for Kant’s conception of self-knowledge.

### B. The Object of Inner Sense

The first of these implications concerns the object of inner sense. Although Kant frequently characterizes this object in traditional terms as the soul, mind, or self, no such object is encountered in inner experience.<sup>3</sup> Moreover, this is a direct consequence of Kant’s denial that inner sense has a manifold of its own. For the latter means that inner sense does not have any data that can be regarded as representations of the soul in the way in which outer intuitions are regarded as representations of body. In addition, Kant also rules out what would appear to be the most obvious candidates for representations of the soul and its states, namely, feelings. Although they may be cognized as objects of inner sense, feelings are denied any representative function, even with respect to the self.<sup>4</sup>

As a result, Kant’s account of inner sense is very close to Hume’s. Just as Hume denied that there is any distinct impression of the self, claiming instead that “when I enter most intimately into what I call *myself*, I always stumble upon some particular impression. . . . I can never catch *myself* at any time without an impression, and can never observe any thing but the impression,”<sup>5</sup> for Kant the only data introspectively available for cognition of the self are its representations, which all derive from outer sense. Nevertheless, there are two essential differences between the two thinkers on this issue. First, Kant, unlike Hume, distinguishes between inner sense and apperception. Thus, the failure to find an inner intuition of the I does not lead to a rejection of the *thought* of the I. Second, and more relevant to our present concerns, Kant attempts to account for a genuinely inner experience.

As Kant describes such experience in what is called the “Leningrad *Reflexion* on Inner Sense,” it occurs insofar as “I bring the representations of outer sense into an empirical consciousness of my state” (LR 10–11). This suggests that inner experience involves a kind of reflective reappropriation of the contents of outer experience. Its content consists of the very representations through which we cognize external objects; but rather than cognizing objects *through* these representations by bringing them under the categories, it makes these represen-



tations themselves into (subjective) objects, which it cognizes as the contents of mental states.

This leads to a significant asymmetry between outer and inner experience, which, in turn, tends to undermine the neat parallelism between outer and inner sense that Kant's account in the *Aesthetic* suggests. In judgments of outer experience, and, therefore, in outer experience itself, representations are taken as representations of the object and are predicated of it in the judgment. But the object itself is viewed as something determinate, already taken under some description, for example, a table, which receives further determination in the judgment, for example, as being rectangular and supporting my computer, rather than as being a "bare particular" or "substratum" to which properties are attached.<sup>6</sup> To be sure, Kant does frequently refer to the object as a "something in general = x," but this characterization applies only to the transcendental object or, perhaps better, to the object considered transcendently, not to the object qua object of outer experience.

Since inner sense has no manifold of its own, there are no sensible representations through which the self can represent itself to itself as object. Consequently, in referring its representations to itself in judgments of inner sense, it does not conceive of them as representations *of itself* in the way outer intuitions are regarded as representations of outer objects. Instead, it conceives of these representations as *belonging to itself*, as its own "subjective objects." Correlatively, the self regards itself merely as the substratum or subject in which these representations inhere. As Kant puts it an important *Reflexion*:

All inner experience is (has) a judgment in which the predicate is empirical and the subject is I. Independently of experience, therefore, there remains merely the I for rational psychology; for the I is substratum of all empirical judgments. [R5453 18: 186]<sup>7</sup>

In addition to the explicit characterization of the I as substratum of all empirical judgments, the most noteworthy feature of this *Reflexion* is its clear implication that this I is non-empirical. Although of itself this is hardly surprising, given Kant's account of apperception, it has as a consequence that the I cannot cognize itself through the empirical predicates (representations) which it refers to itself in judgments of inner experience. Or, more precisely, it cannot cognize itself in the same way in which it cognizes outer objects through the predicates which it attributes to them in judgments of outer experience. Whereas in the latter case both the predicates (derived from outer sense) and the objects of which they are predicated are empirical, in judgments of inner sense this cannot be said of the object. As non-empirical, the I (soul, mind, or self) is not it-

self an object of inner experience or inner sense.<sup>8</sup> These objects are rather the representations that it attributes to itself as “subjective objects.”

### C. Inner Sense and Transcendental Ideality

As a result of all this, the application of the transcendental distinction to the self becomes extremely problematic. If it is regarded as the substratum or owner of its representations, which seems to be the view to which Kant is committed, then it cannot be said to appear to itself at all. Consequently, in contrast to outer objects, there seems to be no basis for distinguishing between the self (or substratum) as it appears to itself and as it is in itself. Nor does it seem to help matters very much if we take the objects of inner sense and inner experience to be the representations themselves, since, as mental entities, the latter are already ideal in the empirical sense. Either way, then, we seem to be without any basis for distinguishing between such an object as it appears and as it is in itself.

Nevertheless, Kant insists upon the ideality of inner sense and the phenomenality of its object and, therefore, upon the doctrine that we cognize ourselves only as we appear to ourselves. Clearly, the main impetus for this doctrine lies in the contrast between inner sense and apperception as two modes of self-consciousness, with the latter yielding the thought, though not the cognition, of the self. But since the self (as substratum) cannot be said to appear to itself through inner sense, this seems insufficient to ground the application of the transcendental distinction to it. Moreover, though Kant certainly appeals to the contrast between inner sense and apperception, his arguments for the ideality thesis do not turn on their juxtaposition. In fact, Kant suggests two lines of argument for this thesis, which will here be called the “materials” and the “self-affection” arguments. We shall see that, though neither is itself capable of establishing the desired conclusion, the second at least points us in the right direction.

1. *The Materials Argument.* The main statement of this argument in the *Critique* is embedded in an overall argument for the transcendental ideality of both outer and inner sense, which Kant added to the Aesthetic in the second edition. This new argument is presented as a “confirmation” of the main ideality argument, which was examined in chapter 5. It consists in the conjunction of the claim that the content of intuition contains nothing but mere relations and the Leibnizian principle that a “thing [*Sache*] in itself cannot be cognized through mere relations.” Kant first applies this to outer sense, maintaining that

the content of outer sense, consisting as it does of mere relations, can yield a representation of the object only in its relation to the subject, not as it is in itself, independently of this relation. He then continues:

It is exactly the same in the case of inner sense. It is not merely that the representations **of outer sense** make up the proper material with which we occupy our mind, but also the time in which we place these representations, which itself precedes the consciousness of them in experience and grounds the way in which we place them in mind as a formal condition, already contains relations of succession, of coexistence, and of that which is coexistent with succession (of that which persists). [B67]

Although Kant is extremely cryptic, which makes any interpretation hazardous, his language certainly suggests that he is offering two independent arguments. The first, which he interjects only in passing, is the materials argument. Reduced to its essentials, it has the following form: (1) since the materials of outer sense are also the materials of inner sense; and (2) since these representations contain nothing but relations; and (3) since a thing in itself cannot be known through mere relations; (4) it follows that we cannot cognize ourselves as we are in ourselves through inner sense.<sup>9</sup>

So formulated, the argument is obviously inadequate, since it involves two distinct non sequiturs. First, even if we assume that sensible intuition contains only relations and nothing “absolutely inner,” it does not follow that such intuition yields a representation of the object only as it is in relation to the subject and not as it is in itself. Kant appears here to conflate two quite distinct theses about the relational character of what is sensibly intuited: (1) that we can sensibly intuit only the *relational properties* of things (because of the spatiotemporal form of sensible intuition); and (2) that we can sensibly intuit objects *only in their relation to the subject*. In addition, there appears to be a conflation of two senses of ‘thing in itself’, namely, the Leibnizian conception of a simple substance or monad, which serves as the non-sensible ground of relations, but which does not itself contain any relational properties, and the transcendental conception of the thing as it is apart from its epistemic relation to the knowing subject.<sup>10</sup>

Second, even granting that through outer sense we cognize objects only as they appear, it does not follow that inner sense yields a representation of the *self* only as it appears. Nor is anything changed by the introduction of the premise that the materials of inner sense are all derived from outer sense. This is because, as we have already seen, outer intuitions, by definition, are not represen-

tations of the self. In fact, if this argument establishes anything, it is that we cannot cognize ourselves at all, at least not through sensible intuition, not that we can cognize ourselves only as we appear to ourselves.

2. *The Self-Affection Argument.* Although the difficult notion of self-affection is central to Kant's account of self-knowledge, it is only discussed in the *Critique* in two places in the second edition. The first is in connection with the previously cited passage from the Transcendental Aesthetic. The second is in §24 of the Transcendental Deduction. In both places Kant's concern is to link this notion with his doctrine of the transcendental ideality of the objects of inner sense. We have seen that in the first passage Kant speaks of the "placing" or "positing" (*setzen*) of representations in the mind and of time as the "formal condition" of this positing. Later in the same paragraph he explicitly equates this positing first with self-affection and then with apprehension. The basic idea is that the mind must somehow affect itself in the act of apprehending its own contents as they appear in inner sense. From this, taken in connection with the doctrine of the ideality of time, it is inferred that the mind can cognize itself only "as it appears to itself, not as it is" (B69).

The account in the Aesthetic does not provide very much enlightenment about the nature of the act of self-affection, but it does indicate the line of argument through which Kant attempts to connect it with the ideality thesis. It turns on the connection between sensibility and affection. As we have seen, Kant maintains that affection by "external objects" is the source of the matter of empirical intuition and, therefore, of the materials of our cognition. Since the mind can receive these materials only insofar as it is affected, it is to that extent passive, and these materials are subject to its form of receptivity. Generalizing this, Kant appears to argue that *anything* cognized on the basis of affection is cognized only as it appears. But since the mind allegedly must affect itself in order to apprehend its contents, it follows that the cognition the mind has of itself is sensible in nature and concerns only the way in which it appears to itself.

As the above sketch indicates, the crux of Kant's argument is that self-knowledge requires sensible intuition, which is itself a consequence of the fact that it involves self-affection. Clearly, this argument is no stronger than the presumed analogy between self-affection and affection by external objects. Unfortunately, however, the analogy does not seem to be strong enough to bear the weight Kant places upon it. The problem is noted by Paton, who remarks that the function of affection by external objects is to supply the raw materials for cognition, while the function of self-affection is to combine these materials in ac-

cordance with the conditions of time.<sup>11</sup> But the original connection that Kant asserts between affection and sensibility rests upon the conception of affection as the source of sensible data, that is, upon an understanding of it as outer affection. Consequently, it hardly seems to follow that a comparable connection with sensibility must be assigned to self-affection, since the latter is not regarded as an independent source of data.

Kant's position becomes even more problematic if we consider the richer account of self-affection given in §24. There Kant explicitly states that by "self-affection" or, equivalently, the "affection of inner sense," is meant the determination of inner sense by the understanding "under the designation of a **transcendental synthesis of the imagination**" (B153). In other words, it is equivalent to the figurative synthesis. Having already dealt at length with the latter and its transcendental function, it is not necessary to discuss it further here. Nevertheless, two points must be noted that relate directly to our present concern. First, since there is little in common between the influence of objects upon outer sense (outer affection) and the "synthetic influence of the understanding on the inner sense" (B154), the identification of self-affection with the transcendental synthesis serves to accentuate the *disanalogy* between the two modes of affection.

Second, the figurative synthesis is a transcendental condition of *all* experience, not merely of inner experience. As we have seen, the main thrust of the argument of the second part of the B-Deduction is to show that the determination of inner sense by the understanding ("under the designation of a transcendental synthesis of the imagination") is necessary in order to provide a determinate intuition for consciousness. This claim is independent of the issue of whether the intuition is of inner or outer objects because all appearances, as modifications of the mind, belong to inner sense. Consequently, the mere appeal to self-affection, construed as transcendental synthesis, hardly explains how such synthesis could serve as a specific condition of inner experience. In fact, by suggesting that it does, Kant appears to be conflating it with the empirical synthesis of apprehension.

3. *The Ideality Thesis of Another Attempt.* Nevertheless, Kant's position is not as confused as the preceding might suggest. The difficulties are real, but they are mainly attributable to the extremely cryptic way in which Kant presents his doctrine. What is needed is a distinction between two senses of "self-affection": one connected with the transcendental synthesis and serving as a condition of all experience, the other connected with the empirical synthesis of apprehen-

sion and serving as a condition of a specifically inner experience. Although Kant never presents this distinction in so many words, it is implicit in the B-Deduction's characterization of the transcendental synthesis as the "first application [of the understanding to sensibility] (and at the same time the ground of all others) to the objects of the intuition that is possible for us" (B152).<sup>12</sup> As we have seen, this "first application" determines the representation of a single universal time in which all appearances have a determinate location. Accordingly, it functions as a transcendental condition of *all* experience. The importance of this for present purposes is its suggestion that there is also a second application (likewise described as self-affection), which is conditioned by, and yet distinct from, the first in the same way in which the empirical synthesis of apprehension is conditioned by, and distinct from, the transcendental synthesis of the imagination. Presumably, this "second application" would be directly involved in inner experience.

This reading is supported by Kant's subsequent appeal to the phenomenon of attention in a footnote to §25. Although Kant confuses matters by suggesting that attention is merely an empirical illustration of the seemingly paradoxical claim that the mind affects itself, it is clear that the real significance of attention is not that it renders intelligible an otherwise mysterious notion of self-affection, but rather that it indicates the specific kind of self-affection required for the institution of inner experience. The point is that in attending to its representations, the mind makes them into objects represented. Thus, instead of perceiving a house by means of a succession of perceptions, all of which are referred to the house as representations thereof, I take this sequence itself as my object. As a second-order, reflective act, this presupposes a prior outer experience and, therefore, the transcendental synthesis of the imagination (the "first application"). Nevertheless, as a "second application," this act involves an active seeking out by the mind of the representations it endeavors to make into objects of inner sense. It also requires a change of epistemic focus and with it a reconceptualization. Whereas the initial conceptualization is the act whereby the given representations are referred to an object, the second is the act whereby these representations themselves become objects.<sup>13</sup> In the last analysis, then, Kant's claim that self-knowledge requires self-affection boils down to the claim that the mind must reconceptualize its representations in order to grasp them as objects.

This account of self-affection enables us to see how it is involved in the determination of the objects of inner sense. Such objects are products of self-affection in the sense that it is only in and through it that the given contents of

the mind can be represented as an object. It is thus constitutive of inner experience, just as the transcendental synthesis is constitutive of experience in general. Moreover, since time is the form of the appearing of representations in inner sense, it follows that time must also be the form in which the products of its own activity appear to the mind in inner experience. As is the case with spatial formal intuitions, it is not that time is imposed upon the manifold of representations by this self-affection; it is rather that this activity is itself constrained or conditioned by the temporal form of the appearing of these representations in consciousness. In this regard at least, the role which time plays in inner experience is analogous to that which space plays in outer experience. Just as space, the form of outer sense, is the form according to which the mind (through its conceptual activity) represents objects as outer, so time, as the form of inner sense, is the form according to which it represents (through a subsequent conceptual activity) something inner (its own representations) as an object. This means that the objects of inner experience, which qua objects are the products of this conceptual activity, are nonetheless sensibly represented. And, as such, they count as appearances in the transcendental sense.

## II. THE REFUTATION OF IDEALISM

It has been suggested that the basic argument of the Analytic is not to be viewed as directed against a Cartesian-type skepticism regarding our putative knowledge of an external world or the "objectivity of experience." Instead, its prime target is a radical empiricism, which challenges the cognitive significance, and thus the validity, of the *a priori* concepts putatively required for judgments about the world, quite apart from the question of the truth or falsity of these judgments in particular instances.

This is not to say, however, that Kant was unconcerned with the more traditional skeptical challenge. On the contrary, his concern with the latter can be traced back to his early writings and is a recurring, albeit subordinate, theme in his philosophical development.<sup>14</sup> In fact, Kant provides two quite distinct treatments of the problem in the two editions of the *Critique*. The first is located in the Fourth Paralogism and attempts to treat the skeptical idealist as a rational psychologist who regards knowledge of the existence of a substantial self as indubitable but the existence of an external (mind-independent) world as problematic, since it is based on a non-evident inference from the contents of consciousness to its cause (A368–80). As previously indicated, Kant regards this form of idealism (there called "empirical idealism") as inseparably con-

nected with transcendental realism, and his critique must, therefore, be understood in that context.<sup>15</sup> By contrast, in the second edition, where the anti-skeptical argument is pointedly termed “Refutation of Idealism,” it is located in the Postulates of Empirical Thinking in General and takes a radically different form, connecting self-knowledge or inner experience with the genuine experience of objects existing in space distinct from the self. Moreover, Kant sought to develop further and clarify this latter line of argument in a series of *Reflexionen* from the late 1780s and early 1790s, which have been given especial significance by Guyer.<sup>16</sup>

The present discussion will focus on the second-edition version (and its variations in the *Reflexionen*) and is divided into three parts. The first offers some preliminary remarks concerning the location, nature, and goal of the argument. The second provides a step by step analysis of its structure. Finally, the third deals with the highly controversial issue of the relation between Kant’s refutation of idealism and his own transcendental idealism.

### A. Some Preliminaries

As noted above, the Refutation of Idealism is located in the Postulates of Empirical Thinking in General, which are concerned with the empirical sense and function of the modal categories. More specifically, it is attached to the discussion of actuality (*Wirklichkeit*) or the actual, which is defined as “[t]hat which is connected with the material conditions of experience (of sensation)” (B265–66). Although Kant’s own official explanation of this location is somewhat misleading, the connection between actuality, so defined, and the Refutation is crucial to the proper understanding of the latter.

In his discussion of the Postulate, Kant’s main concern is to show that, even though the criterion of actuality involves perception, something need not itself be perceived in order to be deemed actual or existent. All that is required is that it be connected with something perceived according to empirical laws (governed ultimately by the Analogies), which Kant describes as “rules for proving existence mediately.” The stated justification for placing the Refutation in this context is that idealism poses a “powerful objection against these rules” (B274). This suggests that the dispute with idealism is directly concerned with the validity of these rules, or, more precisely, the *a priori* principles on which they are based. But this conflicts with the fact that the validity of these principles has supposedly just been established in the Analogies. Consequently, a more plausible interpretation of Kant’s explanation is that the idealist’s challenge is to the premise underlying the applicability of these rules rather than to the rules



themselves. What the idealist questions is the assumption that being connected with sensation (a mental occurrence) is a reliable mark of actuality with regard to the external world.<sup>17</sup>

Although on this reading the traditional skeptical worry is not viewed as a direct threat to the results of the *Analytic*, since it arises at a different level of reflection, it certainly calls into question the *significance* of these results by raising the fresh specter that our categorially structured experience might be nothing more than a consistent dream. In fact, the very presence of a separate refutation of idealism after the completion of the central argument of the *Analytic* is a clear indication that Kant did not believe that the latter addressed this worry. Kant's awareness of the need for doing so is evidenced by his famous remark in the footnote in the B-Preface devoted to the Refutation that "it always remains a scandal of philosophy and universal human reason that the existence of things outside us . . . should have to be assumed merely on faith" (Bxxxix).

In the same note Kant also tells us that he is providing a "strict proof," indeed the only possible one, of the "objective reality of outer intuition" (Bxxxix).<sup>18</sup> In reality, however, the Refutation is addressed to only one form of "material idealism," which is there Kant's generic term for all forms of idealism opposed to his own "formal variety." Thus, it can hardly be regarded as a self-contained proof (strict or otherwise) of the "objective reality of outer intuition." Specifically, it leaves aside the "dogmatic" form of such idealism, which supposedly declares space to be "something that is impossible in itself" and, therefore, "things in space to be merely imaginary" (B274). Kant identifies the latter with Berkeley and excuses himself from dealing with it here on the grounds that it has already been disposed of in the *Transcendental Aesthetic*.<sup>19</sup>

The remaining and more serious threat is the "problematic idealism" of Descartes, which is the explicit target of the Refutation. What makes it more serious is that its thrust is epistemological rather than metaphysical. Instead of denying the reality of space and of objects contained therein, it denies merely that we can have immediate experience, and therefore certainty, regarding the *existence* of such objects. As Kant characterizes this position, its basic claim is that there is only one indubitably certain empirical assertion, namely, "I am" (B274). Consequently, the existence of anything distinct from the mind and its contents can only be established inferentially, and any such inference is problematic.<sup>20</sup>

Against this skeptical position Kant purports to demonstrate the "theorem" that "[t]he mere but empirically determined consciousness of my own existence proves the **existence of objects in space outside me**" (B275). But, given

the nature of the project, the demonstration or “strict proof” is necessarily indirect, since in order to avoid begging the question, the argument must be based on a premise concerning empirical self-knowledge that the Cartesian skeptic will accept, which means one that shares the indubitability of the “I am.”

Kant’s strategy is to exploit an ambiguity in the Cartesian’s understanding of the latter proposition, which is based on the conflation of the two forms of self-consciousness: apperception and inner sense. On Kant’s analysis, as a direct result of this conflation, the Cartesian appeals to the *cogito* inference to secure the existence of the “I” by focusing merely on the activity of thinking, abstracting completely from the content of thought. But, according to Kant, this yields only the empty thought of the “I” as logical subject of thought, not the cognition of a determinate thinking being. Like all cognition, the latter requires intuition, in this case, inner intuition, which supplies the necessary content. Moreover, this means that even the most rudimentary bit of self-knowledge, say, that I am currently perceiving a white surface, rests on inner sense and its conditions. As Kant puts this familiar point in the Refutation,

Of course, the representation **I am**, which expresses the consciousness that can accompany all thinking, is that which immediately includes the existence of a subject in itself, but not yet any **cognition** of it, thus not empirical cognition, i.e., experience; for to that there belongs, besides the thought of something existing, intuition, and in this case inner intuition. . . . [B277]

We can see from this that Kant’s second-edition critique of Cartesian problematic idealism is inseparable from the account of inner sense. In fact, we shall see in chapter 12 that the charge of conflating apperception and inner sense underlies Kant’s whole critique of what may be called the “Cartesian project,” that is, the attempt to arrive at certainty about the nature and existence of the self as *res cogitans* simply by reflecting upon what must be presupposed as a condition of thinking. This will likewise make it clear that the Refutation of Idealism, the account of inner sense, and the revised version of the Paralogisms are integral parts of this critique and must be understood as such, though they also have broader implications. Our present concern, however, is with the argument of the Refutation itself, which we are now in a position to consider.

## B. The Argument

As presented in the *Critique*, the argument consists of five steps. Following our usual procedure, each will be discussed in turn. Additional material from the *Reflexionen* will be introduced when appropriate.

[Step 1.] I am conscious of my existence as determined in time. [B275]

Since this is the premise that must be accepted by the idealist and supposedly generates the *reductio*, its correct interpretation is clearly a matter of some concern. As is so often the case, however, it is also a source of disagreement. One question, which received some attention in the older literature, is whether this consciousness (empirical self-consciousness) should be identified with a mere awareness or with the actual empirical knowledge of the self.<sup>21</sup> Another is the question of what exactly one is supposed to be conscious (or have knowledge) of when one is conscious of one's existence as determined in time. The first question concerns the epistemic status of this form of consciousness, and the second its content.

To begin with, it seems relatively clear that the kind of consciousness that Kant has in mind involves actual self-knowledge rather than mere awareness. Textual support for this is provided by the fact that Kant identifies this consciousness with inner experience, which consists in a cognition (B275).<sup>22</sup> More important, the logic of the argument requires that the premise be taken in this way, since it must correspond to a *proposition* that the Cartesian places beyond doubt.<sup>23</sup>

The second question is more difficult, since there appear to be at least two plausible answers, both of which have a basis in the Kantian texts. One, which we shall term a "thick" conception of empirical self-knowledge, takes the premise to be that the self has some knowledge of its past mental states. As Bennett puts it, "When Kant speaks of 'the determination in time' of my existence he means the establishment of the empirical facts about me—of what my [mental] states have been—at various stages of my history."<sup>24</sup> Similarly, Bennett glosses Kant's "determinate consciousness of the self" as "a consciousness of the self *as* determined in certain ways, as having this rather than that history."<sup>25</sup> By developing what he terms the "realism argument," Bennett tries to show that this relatively thick premise about self-knowledge leads to an argument against the possibility of a purely inner experience. Very roughly, the basic claim is that self-knowledge requires knowledge of the past and this presupposes knowledge of "outer" events and states of affairs.<sup>26</sup>

In addition to its intrinsic philosophical interest, such a reading seems to be supported by Kant's emphasis in the *Critique* on the determination of one's existence in time, which certainly suggests that he is talking about a subject's knowledge of its own mental history. Moreover, beyond the *Critique*, further support is apparently provided by the way in which Kant distinguishes between

transcendental and empirical apperception in the “Leningrad *Reflexion* on Inner Sense.” As he there puts it, “The first says merely I am, the second I was, I am and I will, i.e., I am a thing of past, present and future time . . .” (LR 23–24).

The alternative reading of the premise in question opts for a relatively “thin” conception of empirical self-knowledge. Rather than assuming that a consciousness of one’s existence as determined in time includes a knowledge of one’s past and, therefore, involves a reliance on memory, it limits this knowledge to the contents of one’s current mental state. In order to fit Kant’s characterization, this consciousness must have a diachronic dimension and consist of at least two successive representations, which are cognized as such. But it need not for that reason be construed as knowledge of one’s previous mental history, except in a minimal sense that does not involve any reliance on memory or, more important, any possibility of error.

Kant’s sketchy account of inner experience, as instituted through self-affectation, is compatible with such a reading. As we have seen, this experience involves a refocusing of one’s attention. Rather than relating representations to objects, these representations and the order of their appearance in empirical consciousness is itself made into a (subjective) object. Thus, in perceiving a house, I might introspectively attend to the fact that I first apprehend the roof, then the front wall, and so on. Considered as a judgment about my *past* mental history, this clearly may be mistaken. Quite apart from the question of whether I was really perceiving a house or merely imagining that I was perceiving one, it remains possible that I may, in fact, have received the representations in a different order. But this possibility does not apply to my present introspective consciousness of having representations in the order *A-B* rather than the reverse. What is crucial here is that this succession (as the content of a single complex thought) is for a single *I think* and the inner experience concerns the succession as it is for this *I think*. It is one thing for me to doubt whether my “recollection” of a past experience is veridical and quite another to doubt that I (Henry Allison) am now “recollecting” such an experience.

The point can also be expressed in more Cartesian terms. Suppose that I have just been created, together with a full complement of “memories” and beliefs about my past existence, by the Cartesian demon or some contemporary analogue thereof. In that case all my judgments about my past would be manifestly false. Nevertheless, it would remain certain that I am conscious of myself as the currently existing subject who has these pseudo-memories and this pseudo-history. This is a “thin” conception of self-knowledge, but as a knowl-

edge of the temporal order of the contents of my present mental state it arguably involves a “consciousness of my existence as determined in time.”

Although it will be necessary to revisit this complex issue, it must suffice for now to note that the main and obvious reason for preferring such a reading is that it enables the Refutation to proceed in a non-question-begging manner. Since it is intended as a *reductio* of the Cartesian position, its success is crucially dependent on finding a premise that the latter will accept. This means that it must be a proposition about self-knowledge, which, on the one hand, has some intuitive content and, on the other, is immune to the hyperbolic doubt emanating from the demon hypothesis and its modern variants.<sup>27</sup>

[Step 2] All determination of time presupposes something **persistent** in perception.

Since Kant presents this as a bare assertion, it is generally assumed that he is appealing to the argument of the First Analogy. This reading is sometimes rejected, however, on the grounds of both a dissatisfaction with the latter and a questioning of its relevance to the present argument. Thus, Bennett remarks pointedly that “the Refutation of Idealism could hold little interest if it presupposed the untruth which is the first Analogy.”<sup>28</sup> This view is echoed by Guyer, who distinguishes three arguments from the First Analogy designed to prove that time-determination requires something permanent. But he finds two of these “unpersuasive” and denies that the third is capable of supporting the conclusion of the Refutation, since it does not foreclose the possibility that the required entity might be simply the self.<sup>29</sup>

Having already considered the argument of the First Analogy, we need not reexamine it here. For present purposes it must suffice to note two features of the efforts of Bennett and Guyer to distance the Refutation of Idealism from it. First, they both seem to assume that if the latter is to be relevant to the former it must somehow of itself ensure that what persists is both spatial and numerically distinct from the self. In other words, they neglect the possibility that its relevance is limited to the present step, which, of itself, obviously does not establish anything about the nature of what persists. Second, they fail to recognize that this step relies only on one portion of the progressive argument of the First Analogy, namely, the backdrop thesis. It will be recalled that this thesis affirms merely the necessity of presupposing some (at least relatively persisting) perceptual surrogate for time itself as a condition of the possibility of determining the temporal relations of phenomena. If there were nothing that endured, then we could not become aware of either the coexistence or the succes-

sion of phenomena in a common, objective time. The present step merely extends this principle to the domain of inner experience. Kant does not need to produce any additional argument here, since this extension is warranted by the generality of the claim of the First Analogy (it refers to *all* succession and co-existence).

**[Step 3.] But this persisting element cannot be an intuition in me. For all the determining grounds of my existence that can be encountered in me are representations, and as such they themselves need something persisting distinct from them, in relation to which their change, and thus my existence in the time in which they change, can be determined. [Bxxxix]**

In the footnote to the Preface dealing with the Refutation, Kant remarks that the above passage is to replace the sentence from the text that reads, "This persistent thing, however, cannot be something in me, since my own existence in time can first be determined only through this persistent thing." Although the reason given for this emendation is the obscurity of the original formulation, a more likely explanation is that Kant realized that the failure to mention intuition left it open for the Cartesian to claim that the thinking subject (*res cogitans*) is itself the required persisting entity, and that the existence of the subject's states can be determined in time by being referred to it. As we have seen repeatedly, Kant acknowledges that we have the thought, though not the intuition, of the I itself as such a subject. Consequently his present point is that precisely because it is a thought and not an intuition, it does not refer to anything determinate which can itself serve to determine the existence of the self and its states in time.<sup>30</sup>

The situation is remedied in the revised version of the premise. Not only is there a reference to intuition, there is also an attempt to explain why the required intuition cannot be "in me." Since all intuitions, qua modifications of inner sense, are "in me" (in the empirical sense), it is clear that the term must here refer to the object intuited (*das Angeschauete*) and not to the intuition (*die Anschauung*) itself. In other words, Kant's claim is that what persists cannot be something inwardly intuited, which means that it cannot be an object of inner sense. Once again, all that we inwardly intuit is the appearing (to ourselves) of our own representations. There is no additional intuition of a subject to which they appear (no impression of the self). Since each of these appearings is a fleeting occurrence, inner intuition does not provide anything capable of determining the existence of the subject in time. Consequently, the persisting entity must be something outwardly intuited, which, given the nature of human sensibility, means something intuited in space.

[Step 4.] Thus the perception of this persistent thing is possible only through a **thing** outside me and not through the mere **representation** of a thing outside me. Consequently, the determination of my existence in time is possible only by means of the existence of actual things that I perceive outside myself.

This is the crux of the argument. The previous steps have established that the cognition of one's inner state rests upon the *representation* of something persisting in space. But if Kant is to refute the skeptic, he cannot rest content with this rather modest conclusion. Instead, he must show that I actually experience or perceive, not merely imagine or believe that I perceive, something persisting, which is precisely what Kant claims in the passage quoted above. To be sure, he is careful to point out in the last of the three scholia attached to the argument that this claim does not entail that *all* my perceptions of outer objects must be veridical. In any given case I could be imagining rather than actually experiencing such objects. The essential point, however, is that whether in a particular instance I am experiencing or merely imagining is itself an empirical question, which can be intelligibly raised only against a presupposed background of the actual experience of persisting things. It is the latter that the skeptic denies and that Kant's argument attempts to establish. According to his own formulation of the project, "Here it had to be proved only that inner experience in general is possible only through outer experience in general" (B278–79).

But has he proven it? Once again, the skeptic could readily accept an entailment relation between beliefs, and thus acknowledge the necessity of outer representations. What he would not grant is the contention that this licenses a conclusion about actual experience or real existence. Moreover, it is clear from the note in the preface that Kant was well aware of this fact. Thus, he remarks:

Against this proof one will perhaps say: I am immediately conscious to myself only of what is in me, i.e., of my **representation** of external things; consequently it still remains undecided whether there is something outside me corresponding to it, or not. [Bxxxix-xl, note]

Arguably the most important task confronting any interpretation of the Refutation of Idealism is to present the Kantian answer to this objection. Unfortunately, however, this is not an easy task. In fact, if we consider the official argument of the *Critique*, the note from the Preface, and the relevant *Reflexionen*, we once again find two distinct lines of argument. Moreover, only one of these provides the basis for an adequate response to the skeptic, and even this one must be developed well beyond the point at which Kant leaves it.

The first, manifestly inadequate, line of argument turns on the nature of

outer sense. Its basic premise, which Kant formulates in the note in the Preface, is that “outer sense is already in itself a relation of intuition to something actual outside me.” Since no one will deny that we at least seem to have an outer sense, that is, that we have representations of outer (spatial) objects, all that is necessary to refute the Cartesian skeptic is to demonstrate the incoherence of the suggestion that we might merely believe or imagine ourselves to have an outer sense. Kant regards this as equivalent to the claim that we have merely an outer imagination but not an outer sense. He expresses his basic response to such a claim in a footnote to the main text:

But it is clear that in order for us even to imagine something as external, i.e., to exhibit it to sense in intuition, we must already have an outer sense, and by this means immediately distinguish the mere receptivity of an outer intuition from the spontaneity that characterizes every imagining. For even merely to imagine an outer sense would itself annihilate the faculty of intuition, which is to be determined through the imagination. [B276–77, note]

This claim seems to be implicit in the actual argument of the *Critique* (in the move from *representation* to *thing* represented), and it reappears explicitly, in slightly different form, in the third scholium attached to the argument. In commenting on the possibility of merely imagining outer things (as in dreams or illusions), Kant remarks there that “this is possible merely through the reproduction of previous outer perceptions, which, as has been shown, are possible only through the reality of outer objects” (B278). In addition, variations of the same basic theme are scattered throughout the *Reflexionen*. Although these formulations differ from one another in detail, they reduce to the assertion of the incapacity of the imagination, either of itself or with the assistance of inner sense, to produce the representation of space or of things in space. Behind this assertion lies Kant’s contention that inner sense has no manifold of its own, from which he infers its incapacity to generate the data of outer sense.<sup>31</sup> Consequently, the mere fact that we have outer representations is taken as proof that we have an outer sense, which, in turn, entails that the mind is affected by, and perceives, actually existing objects (though not as they are in themselves).

It is not necessary to pursue any further the details of this line of argument in order to recognize its inadequacy. For one thing, the contention that we could not even imagine or dream about outer objects unless we had an outer sense is similar to one Descartes himself entertains and rejects in the First Meditation. For another, it rests upon some dubious claims about the capacity (or lack



thereof) of particular faculties to produce particular species of representations. But even if, for the sake of argument, these claims are accepted, the possibility still remains open that our representations of outer things are the results of some unknown “hidden faculty.” It will be recalled that Descartes raises this very possibility in the Third Meditation in connection with his argument for the existence of God. Although Descartes denies that such a faculty could be the source of the Idea of God as a most perfect being, the skeptic could readily respond to Kant at this point by suggesting the possibility that some such faculty is the source of our representations of outer things.<sup>32</sup> In short, because of its dogmatic assumptions, the argument from the nature of outer representations cannot undermine the possibility that the contents of consciousness, that is, the sequence of “outer” representations, might be precisely as they in fact are, without there being anything external to the mind corresponding to them.

The second line of argument, which is encapsulated in the claim that in the preceding proof “the game that idealism plays has with greater justice been turned against it” (B276), seems considerably more promising. Expressed in Kant’s terminology, the idealist’s “game” is to assume the indubitability of inner experience and, using this as a model of perfect certainty, to cast doubt on the reality of outer experience on the grounds that it fails to measure up to this norm. To turn this game against the idealist is to argue that the inner experience, which the latter acknowledges to be non-problematic, presupposes the reality of outer experience. Consequently, the critique is immanent, turning on a premise to which the idealist is supposedly committed and claiming that this commitment entails a commitment to what the idealist regards as questionable. Kant puts the point succinctly in the note added to the B-Preface when he remarks, “I am just as certainly conscious that there are things outside me to which my sensibility relates, as I am conscious that I myself exist as determined in time” (Bxli). And, somewhat more expansively, he writes in a *Reflexion*:

The proposition is: the empirical consciousness of our existence in time is necessarily bound up with the empirical consciousness of a relation to something outside us and the one is just as little an illusion from a mistaken inference, even just as little an inference, as the other. [R5653 18: 308]

In these passages Kant asserts that inner and outer experience possess the same epistemic status. A similar thesis was already affirmed in the first edition, where Kant insisted that they are both “immediate” in the sense that neither involves any inference from representation to external object as cause (A370–71).

But now Kant goes further, arguing that outer experience “wears the trousers,” in the sense that it conditions inner experience. Consequently, the price of doubting outer experience is to render inner experience similarly dubitable.

Although the orthodox Cartesian might be unwilling to pay this price, a more radical form of skepticism (one concerning self-knowledge) might not. Indeed, given the “thick” conception of self-knowledge advocated by Bennett, such a skepticism initially seems plausible. Nevertheless, the “thin” conception characterized earlier is immune from such doubts. If I am conscious of a succession of representations, then, by that very fact, these representations really do succeed one another in my consciousness. In other words, though there can be a succession of representations without a consciousness of their succession, the converse does not hold. My consciousness of such a succession is at the same time a succession in my consciousness. Moreover, it is precisely the latter that is claimed to be known with certainty.

Still, this does not complete the argument, since the skeptic might question the equal-certainty principle. Specifically, it might be asked why the outer experience that is presumed to be a condition of inner experience could not be itself something illusory. If, as has been admitted, Kant’s claims about the possible source of our outer representations are insufficient to answer the skeptic, then it does not seem that the argument up to this point suffices to remove this possibility.<sup>33</sup> Thus, the question becomes whether the consciousness of one’s existence as determined in time would lose its epistemic status, if it were conditioned by something illusory. Although Kant may not have had precisely this worry in mind, the fifth step of the argument can be viewed as an attempt to respond to it.

[Step 5.] Now consciousness in time is necessarily combined with consciousness of the possibility of this time-determination: Therefore it is also necessarily combined with the existence of the things outside me, as the condition of the time-determination; i.e., the consciousness of my own existence is at the same time an immediate consciousness of the existence of other things outside me.

Here Kant makes fully explicit the dependence of inner on outer experience that has been implicit all along. The same thought is also expressed more emphatically in a passage from the note added to the Preface where Kant states:

This consciousness of my existence in time is thus bound up identically with the consciousness of a relation to something outside me, and so it is experience and not fiction, sense and not imagination, that inseparably joins the outer with my inner sense; for outer sense is already in itself a relation of intuition to something actual

outside me; and its reality, as distinct from imagination, rests only on the fact that it is inseparably bound up with inner experience itself as a condition of its possibility, which happens here. [Bxl, note]

And later in the same note he adds:

The representation of something **persisting** in existence is not the same as a **persisting representation**; for that can be quite variable and changeable, as all our representations are, even the representations of matter, while still being related to something permanent, which must therefore be a thing distinct from all my representations and external, the existence of which is necessarily included in the determination of my own existence, with which it constitutes only a single experience, which could not take place even as inner if it were not simultaneously (in part) outer. [Bxli, note]

The language of being “combined” or “bound up identically with” (*identisch verbunden mit*) indicates the inseparability of inner from outer experience. In the second of the passages quoted from the note, Kant attempts to explicate this inseparability by suggesting that inner and outer experience constitute two facets of a single experience, and even that inner experience is also “(in part) outer.” There appear to be two factors underlying this claim. One is that the objects of both modes of experience exist in a single universal time and are subject to its conditions. The other is the peculiar nature of inner experience: it does not contain any inner persisting object, which means that the object required for the determination of time must be found without. Thus, to claim that inner experience is in part outer is to claim that it requires an outwardly intuited persistent as a condition of its own possibility. The converse does not apply, however, since outer experience, having its own persisting object, does not require inner experience as a condition of its possibility. Consequently, even though time is the form of inner sense and cannot be outwardly intuited, one cannot say (at least not in the same sense) that outer experience is in part inner.

Although none of this is really new, it is essential for understanding and evaluating Kant’s argument. In particular, it provides the basis for rejecting the possibility that the consciousness of one’s existence in time could be determined, if there were not something genuinely outer that persists on the basis of which this inner existence is determined. Put simply, if inner experience is really inseparable from outer experience, if they are two facets of a single experience, and if all inner experience is also in part outer, then inner experience cannot be veridical and the outer experience that conditions it illusory. What would make outer experience illusory, namely, the nonexistence of the things that I perceive

to persist through change and to be distinct from myself, would thereby do the same for my putative inner experience. Thus, I might *imagine* that I am conscious of my existence as determined in time but I would not really be so, since, *ex hypothesi*, there would be nothing capable of determining it.

Admittedly, this is more of a sketch for a possible refutation than a fully worked out argument, capable of answering the Cartesian skeptic. Considerable development and spelling out of details are necessary, if it is to become an adequate refutation. For example, in some of his *Reflexionen* Kant hints at an argument showing that one's body functions as the persisting object with respect to which one's existence is determined in time.<sup>34</sup> Presumably, the temporal order of one's mental states is determined by their correlation with one's bodily states, and through this connection one's existence is determined with respect to that of other objects in the "field of experience." There is no need to insist upon any of this here, however, since the important point is only that, even in its embryonic form, this second line of argument is both distinct from and superior to the first. Whereas the former involves a questionable claim about what must be presupposed in order to account for some of the representations that we do possess, this one contends that a mode of experience that is accepted by the skeptic is conditioned by, and inseparable from, another mode, which the skeptic does not accept. It thus provides at least the beginnings of a genuine *reductio* of the skeptic's position.

### C. The Refutation of Idealism and Transcendental Idealism

One of the persisting worries concerning Kant's refutations of idealism is their compatibility with his transcendental idealism. The worry arises in a particularly sharp form in relation to the first-edition refutation, where it has seemed to many that Kant only succeeds in dismissing the problematic idealism of Descartes by the devious device of reducing supposedly external objects to "mere representations," thereby adopting a phenomenalist position or "subjective idealism" that is in essential agreement with Berkeley.

It was argued in chapter 2 that the latter involves a serious misrepresentation of Kant's position, since even in the first edition his idealism is quite distinct from any of the usual forms of phenomenism and he does not endeavor to refute Descartes by siding with Berkeley. Even granting this, however, it might be wondered whether transcendental idealism, correctly understood, is at all relevant to the second-edition version of the Refutation and its variants in the *Reflexionen*.

Not surprisingly, this line of questioning has been pursued vigorously by Guyer, among others.<sup>35</sup> In fact, Guyer insists that if the Refutation is to have any hope of success, the persisting objects with respect to which one's inner state is determined in time must be accorded more than merely phenomenal status. Moreover, he maintains that in his later *Reflexionen*, if not in the published Refutation, Kant himself embraces just such a view. Indeed, for Guyer this is an essential ingredient in his larger thesis that whatever is philosophically viable in Kant is independent of transcendental idealism.

Actually, Guyer's effort to divorce the Refutation of Idealism from Kant's transcendental idealism has undergone an interesting development. A constant in Guyer's reading is the curt dismissal of the A-version on the grounds that it contains a phenomenalistic form of idealism that is incapable of yielding a genuine refutation.<sup>36</sup> He also consistently insists that this problem is avoided in the successful strategy in the later versions, since Kant there argues first for the ontological independence of the persisting object required to "interpret" and "justify" the order of one's representations, and only then infers its spatiality on the grounds that space is the form of outer sense.<sup>37</sup> He thus grudgingly admits that the Refutation is compatible with a non-phenomenalistic form of transcendental idealism to which he believes Kant subscribed in 1787 and after (though not in 1781). According to this form, objects exist ontologically (not merely phenomenologically) distinct from our representations of them, though because of our necessary reliance on our forms of sensibility we cannot cognize them as they are in themselves. It is not, of course, that Guyer endorses this form of idealism, merely that he admits it is not precluded by the later versions of the Refutation.

More recently, however, Guyer seems to have gone beyond this, maintaining that, even though the Refutation may be compatible with the ideality of space, it is not compatible with that of time. Appealing to the earliest critics of this doctrine (Lambert, Mendelssohn, and Sulzer), Guyer argues that "Kant cannot consider duration . . . a feature of our representations which is not also a real feature of objects distinct from our representations," because "his theory of time-determination depends precisely on the assumption that objects other than representations *"do endure"* in a way that our 'transitory' and 'changeable' representations do not."<sup>38</sup> And from this it presumably follows that such objects must be things in themselves in the transcendental, not merely the empirical, sense.

This analysis is, however, far from compelling. To begin with, it seems to turn on an ambiguity in the phrase "*endure* in a way that our 'transitory' and 'changeable' representations do not." Taken in one sense this is undeniably cor-

rect, since the argument clearly requires that objects *endure*, that is, persist throughout a period of time, whereas our representations do not. But it does not follow from this either that the time in which objects endure has an ontological status different from the time of our transitory and changeable representations or that the time of both must be something other than transcendently ideal. On the contrary, there is no reason why the transcendental ideality of time is not compatible with its empirical reality in the same way as the ideality of space is compatible with its empirical reality. As should be clear by now, it is not as if this would commit Kant to the view that objects only *seem* to endure, whereas they really do not. These objects *really* persist in precisely the same way as our representations *really* change. It is only that neither this persistence nor this change has any significance, if we consider things in abstraction from our forms of sensibility.

Moreover, the Refutation of Idealism is not merely compatible with transcendental idealism, properly construed; it presupposes it. In order to appreciate this we must keep in mind that its goal is to demonstrate the objective reality of outer intuition, that is, the existence of objects in space (Bxxxix). In the Refutation itself Kant characterizes the skeptical idealist as concerned with the reality of outer experience. Thus, in order to defeat such an idealist it is necessary to show that outer experience as such (as distinct from particular “experiences”) cannot be dismissed as illusory, since it is the condition of the inner experience that the idealist accepts. But this goal cannot be accomplished on the transcendently realistic assumption that our outer intuition or experience must be of things as they are in themselves, because that assumption opens up an epistemological gap between the representation of something outer and its purported object, which the skeptic is able to exploit. As Kant himself puts it in a *Reflexion*,

If our cognition of outer objects must be a cognition of them (and of space) as things in themselves, then we would never be able to prove the reality of such objects from our sensible representation of them (as external to us). For only representations are given to us, the cause of which can be either in us or outside of us, regarding which the senses decide nothing. But if the representations of inner sense as well as outer are merely representations of things in the appearance and even the determination of our consciousness in inner sense is possible only through representations outside of us in space. . . [R6313 18: 614–15]

This text is noteworthy because it contains an interesting combination of motifs from the first- and second-edition versions of the Refutation, thereby

indicating that Kant never fully abandoned the argument of the former. More specifically, Kant is here apparently attempting to combine the first-edition version's core thought, that if the objects of cognition are taken in the transcendently realistic sense as things in themselves we could never have epistemic access to them, with the second-edition version's essential move that the inner experience, which the skeptic allows, presupposes the outer experience that is deemed problematic. Thus, even though the passage breaks off in mid-sentence, we can easily complete it on the basis of the second-edition argument. The point is simply that on the transcendently idealistic assumption that we experience things as they appear to us in accordance with our forms of sensibility ("are merely representations of things in appearance"), we can readily see why the determination of our consciousness in inner sense (inner experience) is only possible on the basis of the experience of persisting objects of space. In other words, Kant continued to hold that the central worry, posed already in the note devoted to the Refutation in the B-Preface, namely, that I am immediately consciousness only of a representation of external things, from which it remains undecided whether there is something outside me corresponding to it (Bxxxix), can only be addressed within a transcendently idealistic framework.<sup>39</sup>

Granted, Kant often insists that the objects we represent as in space and time must also be thought as existing in themselves, which is the point on which Guyer focuses. Nevertheless, this should not be viewed as the main thrust of the Refutation, at least not on the basis of what has here been characterized as the successful strategy, which is based on a *reductio* of the skeptical idealist's claim that the only secure form of experience is the inner variety. Since this challenge is to the reality of outer *experience*, that is, our purported cognition of persisting objects in space, it is not sufficient to respond to it by appealing to things that exist in themselves, when, *ex hypothesi*, they are not experienced as they are in themselves. In other words, once the transcendental distinction between two ways of considering things is introduced, as it must be in order to evade the epistemological gap exploited by the skeptical idealist, it will not do to appeal to unexperienceable things as they are in themselves to secure the validity of experience. On the contrary, it is only if we assume that we cognize things as they appear to us in virtue of the forms of our sensibility rather than as they are in themselves that we have any hope of establishing the reality of things that are ontologically distinct from the self.

Moreover, Kant makes precisely this claim in a later *Reflexion* (dated 1790 or 1791). As he there puts it,

If, however, it is shown [against the idealist] that the determination of our own existence in time presupposes the representation of space, in order also to be able to represent to oneself the relation of the determinations of inner intuition to the persisting object . . . then reality can be secured for outer objects (as things in themselves) [*als Sachen an sich*] precisely insofar as one does not assume that their intuition is one of a thing in itself [*einer Sache an sich*]; for if it were this, and the form of space were the form of a thing which pertained to it even without the special constitution of our subject, then it would be possible that we should have the representation of a thing without it also existing. But it is a special kind of intuition in us, which cannot represent what is in us, thus what exists in the flux of time, because then as a mere representation it would be capable of being thought only in relations of time; therefore such an intuition must stand in a real relation to an object outside us and space actually signifies something, which is possible to represent in this form of intuition only by means of a relation to an actual thing outside us. [R6317 18: 627–28]<sup>40</sup>

Although Kant appears to be claiming that in order to defeat the skeptical idealist it is necessary to establish the reality of outer objects as things in themselves rather than as mere phenomena, he also insists that this can be accomplished only insofar as we are prepared to deny that our representation of these objects as in space is an intuition of them as they are in themselves. As Kant proceeds to make clear, this is once again because of the epistemological gap between representation and object that arises on the assumption that our outer intuition must correspond to things as they are in themselves. But, Kant continues, this problem is avoided if we take the representation of space to be “a special kind of intuition in us,” since (as outer) “such an intuition must stand in a real relation to an object outside us.”

Admittedly, the latter part of this claim is dubious; indeed, it is an instance of the first of the two strategies distinguished earlier. Nevertheless, even though it may not correspond precisely with what Kant here had in mind, the point can be reformulated in terms of the successful *reductio* strategy. So construed, the move is from the actuality of outer experience (regarded as a condition of the possibility of the inner experience that the skeptical idealist accepts as non-problematic) to the existence of objects that we cannot cognize as they are in themselves. In other words, the genuineness of outer experience entails the actuality of the phenomenal objects of outer sense, which, in turn, entails the reality of something that exists in itself. For the object perceived as persisting is something that exists in itself (as does every object of experience), though it is not cognized as it is in itself. More important, however, it is precisely because it is *not cognized* as it is in itself that it is able to serve as a condition of time-deter-



mination. Consequently, while Guyer is correct in pointing out that the outer objects we experience under the conditions imposed by our sensibility must be assumed to have an existence in themselves that is ontologically distinct from the self, it remains the case that the B-version of the Refutation is dependent on transcendental idealism.

## Part Four **The Transcendental Dialectic**

## Chapter 11 Reason and Illusion

Although reason is assigned the title role in the *Critique of Pure Reason*, its first significant appearance is in the third act, the Transcendental Dialectic. To be sure, reason made a cameo appearance at the beginning of the Transcendental Logic (A63 / B88); but there it is not distinguished from the pure understanding, the non-empirical use of which is identified as the source of dialectical illusion.<sup>1</sup> Moreover, the lengthy critique of the pretensions of reason in the Dialectic is frequently viewed as largely redundant, since much of the demolition of traditional metaphysics was already accomplished in the Analytic through the limitation of knowledge to possible experience. In fact, in the very last paragraph of the Dialectic, Kant admits that the Analytic, of itself, suffices to establish the limitation thesis and suggests that were it not for a continuing illusion connected with metaphysical claims, we might have been spared the additional labors of the Dialectic (A702–03 / B730–31).

Nevertheless, in the same context Kant also points out that the illusion persists even after the demonstration of the untenability of all claims of knowledge of objects beyond those of possible experience.

Indeed, his view is that the illusion of such an extension of knowledge is both natural and unavoidable. Consequently, a transcendental critique has the two-fold task of removing its deceptiveness, by exposing its source in the nature of human reason, and demonstrating the fallacious nature of the inferences drawn under its influences regarding the soul, the world, and God, the putative objects of pure reason, in the metaphysical “sciences” of rational psychology, rational cosmology, and rational theology, whose provenances are these objects.<sup>2</sup>

The Dialectic contains, however, more than a systematic exposure of the underlying illusion and the specific fallacies flowing from it. This exposure is grounded in an analysis of reason as a cognitive faculty distinct from, yet closely related to, the understanding. Moreover, in spite of the less than sympathetic reception that Kant’s account of the theoretical function of reason has received, it will be argued here that it is both of considerable philosophical interest and indispensable for a proper understanding of the *Critique*. This, then, is the subject of the present chapter, which is intended as an introduction to the fourth part of this study and is divided into two parts. The first provides a sketch of Kant’s conception of theoretical reason and its essential products, the transcendental ideas; the second analyzes the peculiar species of illusion (transcendental illusion) that is supposedly inseparable from its use.<sup>3</sup>

## I. REASON

In contrast to the treatment of the understanding at the beginning of the *Analytic*, where he was content to distinguish between its logical and its transcendental uses, Kant commences his account of reason with a brief consideration of “reason in general” or “as such” (*überhaupt*) and only then distinguishes between its logical and its real (transcendental) uses. The significance of this extra layer of complexity lies in the fact that it indicates that, contrary to the widespread assumption, Kant did not simply tack his account of transcendental ideas onto a traditional view of syllogistic reasoning (the logical use of reason). On the contrary, Kant notes that the traditional logical conception of reason as the faculty of drawing mediate inferences provides no insight into the nature of reason considered as a transcendental faculty with a real, that is, metaphysical, use. Moreover, he suggests that both its logical or syllogistic and its real use (based on ideas) must be seen as expressions of a generic function of reason (A299 / B355–56). Accordingly, we shall begin by considering the latter, then turn to the question of its real use, and finally proceed to an examination of the transcendental ideas and their “metaphysical deduction.”

## A. Reason in General

Kant begins his generic account of reason by remarking, “All our cognition starts from the senses, goes from there to the understanding, and ends with reason, beyond which there is nothing higher in us to work on the matter of intuition and bring it under the highest unity of thinking” (A298 / B355). Clearly, this progression is logical rather than temporal. Just as the function of the understanding is to unify the raw material given in sensible intuition by bringing it to the objective unity of apperception, the function of reason is to unify the discrete products of the understanding (judgments) by bringing them into a coherent whole (a system). The work of reason thus stands at the apex of the cognitive enterprise; if attained, the unity at which it aims would constitute the completion of knowledge.



Kant endeavors to explicate reason’s supreme status by characterizing it as the **“faculty of principles [Prinzipien],”** in contrast to the understanding, which had previously been defined as the “faculty of rules” (A299 / B356). But since the difference between rules and principles is not immediately apparent, Kant finds it necessary to clarify the matter by distinguishing between a broad and a strict sense of the latter.

Taken in the former sense, any universal proposition from which consequences can be deduced counts as a “principle.” Since this includes products of the understanding, it does not suffice to differentiate the two faculties. Indeed, Kant reminds us that he had already spoken in the *Analytic* of “Principles [Grundsätze] of Pure Understanding.” These principles, however, yield synthetic cognition only in relation to sensible intuition (or possible experience). Consequently, principles in the strict or “absolute” sense (if there be any) would yield synthetic cognition from concepts alone, without any appeal to intuition (A301 / B357–58). It is such principles that Kant connects with the use of reason.

For readers of the *Analytic*, this seems deeply mysterious, since one of its major lessons is that synthetic cognition requires intuition as well as concepts.<sup>4</sup> Kant is not claiming, however, that reason can provide such cognition. On the contrary, he notes that the very idea of such cognition, “if not impossible, is at least very paradoxical in what it demands” (A302 / B358). Nevertheless, we shall see that it is necessary to recognize that reason, considered as distinct from the understanding, does involve such a demand in order to comprehend both how it gives rise to a natural and unavoidable illusion and how, liberated from the deceptiveness of this illusion, it yields regulative principles that are essential for the progress of the understanding within experience.


In this introductory discussion, however, Kant's purpose is to contrast the kind of unity aimed at by reason in its projected knowledge from principles with that produced by the understanding through rules and to clarify the connection between them. As he puts it, "If the understanding may be a faculty of the unity of appearances by means of rules, then reason is the faculty of the unity of the rules of understanding under principles" (A302 / B359). In other words, whereas the understanding is a first-order faculty, which relates directly to intuition, reason is a second-order faculty, whose concern is with the conceptually determined products of the former. As Kant points out, this makes the unity of reason qualitatively different from the unity produced by the understanding.

Although Kant does not here inform us about the cognitive significance of this "highest unity of thinking," the essential point is that it is required to complete the fragmentary knowledge attained by the understanding. Basically, what the understanding endeavors to do is to explain appearances by bringing them under rules. Thus, its principles provide the transcendental conditions of the possibility of doing this. But without delving too deeply into the topic, it should be clear that the explanation of any given appearance depends crucially on connecting it with others (as a token of a type). Similarly, the explanation of a type, say gases, requires understanding their characteristic behavior as falling under a more general set of explanatory rules or laws, say, those referring to the molecular level, and so forth.

The picture, then, is of the understanding, in its quest to unify phenomena, being driven by the inherent logic of explanation to ever more comprehensive sets of rules. The ultimate goal of this endeavor is what Kant terms the "unity of reason," which amounts to the interconnection of the dispersed, fragmentary cognitions of the understanding under a single principle. As Kant points out in the Appendix to the Dialectic, this is a projected unity, and it rests on the idea of "the form of the whole of cognition, which precedes the determinate cognition of the parts and contains the conditions for determining *a priori* the place of each part and its relation to the others" (A645 / B673). In other words, the discursive human understanding is oriented from the beginning toward the idea of such a whole, which could alone provide the requisite closure. Thus, even though it remains the understanding that does the unifying, the ideal, projected unity at which it aims differs in kind from any unity attainable by the understanding alone and is the distinctive contribution of reason.<sup>5</sup>



## B. The Real Use of Reason

After a brief discussion of the “logical use of reason” or syllogistics, Kant turns to the question of whether reason has a real or, as he terms it here, “pure” use. In other words, the question is whether the function of reason is merely to organize already given cognitions into convenient parcels, the members of which are related by logical implication, or whether it also serves as an independent source of knowledge. In view of the preceding definition of reason, this is equivalent to the question of whether reason has any principles in the strict sense that are capable of yielding synthetic *a priori* knowledge independently of any reference to intuition or the possibility of experience. 

In addressing this question, Kant focuses on the principle operative in reason’s direction of the understanding, and, as one might expect, the logical (syllogistic) function of reason provides the clue to the discovery of the transcendental principle unique to reason. This function is to infer or, more precisely, to draw mediate inferences.<sup>6</sup> What makes an inference “mediate” in Kant’s sense is not simply the number of premises, but that it involves the subsumption of a judgment, taken as “conditioned,” under another judgment, taken as “condition,” by means of an additional, intervening condition. Moreover, it is just this procedure that is at work in the syllogism or inference of reason [*Vernunftschluss*], which Kant here characterizes as “nothing but a judgment mediated by the subsumption of its condition under a universal rule (the major premise)” (A307 / B364).<sup>7</sup>

This logical procedure is illustrated by the paradigmatic categorical syllogism: All humans are mortal/Caius is human/Caius is mortal. Here the rule affirmed in the major premise is that all humans are mortal and “the condition of the rule,” that is, the condition under which the rule applies, is that one is human. Given this, the minor premise affirms that Caius falls under that condition, which, in turn, licenses the inference: Caius is mortal. Generalizing from this, Kant elsewhere suggests that the universal principle governing all inferences of reason may be expressed in the formula “*What stands under the condition of a rule also stands under the rule itself*” (JL 9: 120; 615).<sup>8</sup>

For understanding the function of reason, however, the crucial point is that the unifying process does not cease with the subsumption of a given cognition under a universal rule by means of a mediating premise. Rather, the rule itself requires its rational grounding, which it can receive only by being derived from a higher principle, and so forth. Thus, reason in its logical use endeavors to order the data provided by the understanding in a series of prosyllogisms. More-

over, in doing so it is guided by the directive “[F]ind the unconditioned for the conditioned cognitions of the understanding, with which its unity will be completed” (A307 / B364).

Kant characterizes the latter as a “logical maxim.” But since it places an unconditional demand on the understanding that is analogous to the one pure practical reason places on the will, namely, a demand of maximal coherence, it might also be regarded as an intellectual categorical imperative. So construed, however, it does not amount to a principle of reason in the required sense, since it neither has propositional form nor can itself give rise to any knowledge. Nevertheless, Kant suggests that it becomes such once we assume that “when the conditioned is given, then so is the whole series of conditions subordinated one to the other, which is itself also given (i.e., contained in the object and its connection)” (A307–08 / B364).

The latter, then, is *the* principle of reason, and we shall see it is also what Kant understands by transcendental illusion. In fact, Kant indicates as much near the end of his discussion, when he raises the possibility that this “need of reason” (for conditions) “has, through misunderstanding, been taken for a transcendental principle of reason, which overhastily postulates such an unlimited completeness in the series of conditions in the objects themselves” (A309 / B355–56). Rather than prejudging the issue of its illusory nature, however, here Kant is content to point out that this principle is synthetic, since “the conditioned is analytically related to some condition, but not to the unconditioned” (A308 / B364). In other words, this principle (in contrast to the logical maxim) involves a metaphysical assumption concerning the reality of a complete set of conditions for every conditioned, a set that must be considered as unconditioned, since, *ex hypothesi*, there can be nothing further by which it is itself conditioned.

Moreover, this principle, which Kant now characterizes as the “supreme principle of pure reason” (A308 / B365), is itself the source of several subordinate principles, all of which, like the supreme principle itself, are transcendent, since the unconditioned objects to which they point can never be met with in a possible experience. These principles govern the regresses from distinct types of conditions to their unconditioned ground. And since in each case the unconditioned must be conceived in terms of the type of condition it grounds, they give rise to a distinct set of concepts unique to reason, which Kant terms “transcendental ideas.”



### C. The Transcendental Ideas as Concepts of Reason


Kant famously begins his discussion of ideas by referring approvingly to the original Platonic understanding of the term.<sup>9</sup> As he points out, Plato used it to characterize objects (of thought) that transcend all possible experience and that serve as archetypes. Consequently, ideas are to be sharply distinguished both from concepts of the understanding or categories, with which, Kant suggests, Aristotle was concerned (A317–18 / B373–74), and from “ideas” in the Cartesian-Lockean sense, which include all representations (A319 / B376).

Less frequently noted, but of equal import for the understanding of Kant’s theory of ideas, is the distinction between reflected and inferred concepts with which he begins the first book of the *Transcendental Dialectic* (A310 / B366). This distinction refers to the origin of the concepts. We have seen that Kant regards all general concepts as analytical unities, consisting of a set of common marks, and as products of the “logical acts” of “comparison,” “reflection” and “abstraction.” We have further seen that a similar procedure is supposed to account for the genesis of the categories, considered as full-fledged concepts capable of serving as predicates in possible judgments. For these too are acquired through an act of reflection rather than being innate, though in this case the reflection is directed to their schemata, functioning as rules guiding the figurative synthesis.<sup>10</sup>

Although transcendental ideas or concepts of reason have a quite different genealogy, which Kant indicates by characterizing them as “inferred,” they are likewise acquired rather than innate.<sup>11</sup> The inference in question proceeds under the direction of the above-mentioned “supreme principle of reason,” and is from some cognition, viewed as conditioned, to the totality of its conditions and, therefore, to something unconditioned that provides the requisite closure. In short, transcendental ideas arise from an inferential process that expresses the inherent dynamic of reason, and they characterize the various ways in which the culmination of this process (the unconditioned) can be conceived.


This suggests that the inference forms involved in the regress from a conditioned cognition to its conditions will provide the “clue” to the discovery of the transcendental ideas just as the judgment forms catalogued in general logic provided the clue to the discovery of the pure concepts of the understanding. Moreover, since these inference forms are to be found in the various forms of syllogism, it is to these we must look in order to find the source of these ideas. As Kant puts it at the beginning of his discussion of the transcendental ideas,



The transcendental analytic gave us an example of how the mere logic  of our cognition can contain the origin of pure concepts *a priori*, which represent objects prior to all experience, or rather which indicate the synthetic unity that alone makes possible an empirical cognition of objects. The form of judgments (transformed into a concept of the synthesis of intuitions) brought forth categories that direct all the use of the understanding in experience. In the same way [*Ebenso*], we can expect that the form of the syllogisms, if applied to the synthetic unity of intuitions under the authority of the categories, will contain the origin of special concepts *a priori* that we may call pure concepts of reason or **transcendental ideas**, and they will determine the use of the understanding according to principles in the whole of an entire experience. [A321 / B377–78]

Perhaps no other doctrine in the *Critique* has met with such a uniformly hostile reception. Indeed, even such a usually sympathetic critic as Walsh expresses a general agreement with Bennett regarding the worthlessness of Kant's entire theory of reason (including its distinction from the understanding).<sup>12</sup> Walsh reserves special scorn, however, for what he terms "Kant's unlikely story about syllogisms," remarking that "[o]ne can only avert one's eyes from this particular folly."<sup>13</sup>

The "folly" supposedly lies in appealing to the syllogism to locate the basic metaphysical concepts, soul, world, and God. Such a criticism amounts to a replay at the level of reason of the dismissive treatment accorded the Metaphysical Deduction, and it rests on a similar distortion of Kant's view. As we saw in chapter 6, Strawson, Guyer, and others reject the Metaphysical Deduction because they assume that Kant was trying to show that we require a particular category in order to be able to judge under the corresponding logical form. The present assumption is that Kant maintains that syllogizing in the categorical manner somehow involves the idea of the soul, in the hypothetical the idea of the world, and in the disjunctive the idea of God. But since any such thesis is patently absurd, Kant's whole account is dismissed as an embarrassment, unworthy of serious consideration.<sup>14</sup>

This line of criticism is misguided here, however, just as it was in connection with the derivation of the categories. With regard to the latter, we have seen (chapter 6) that Kant's actual view is that logical function and category express one and the same function of the understanding operating at two levels: as a rule for the unification of concepts in a judgment and as a rule for the synthesis of intuitions in experience. Similarly, in the case of reason, it is one and the same function (seeking the totality of conditions for a given conditioned) that is operative in both syllogistic reasoning and the metaphysical reasoning lead- 

ing to the transcendental ideas as distinct concepts of the unconditioned.<sup>15</sup> Although this claim may be far from obvious, it is not patently absurd. Accordingly, it is worth considering why Kant thought it to be true.


As is often the case, Kant's official explanation is not immediately helpful. It begins with a brief reconsideration of what has supposedly already been established, namely, that "[t]he function of reason in its inferences consisted in the universality of cognition according to concepts, and the syllogism is itself a judgment determined *a priori* in the whole domain of its condition" (A321–22 / B378). This characterization of the inferential function of reason is rather opaque, but Kant attempts to clarify it by means of a brief analysis of the previously considered categorical syllogism. In accordance with his emphasis on the regressive side of the inferential process, Kant begins with its conclusion, "Caius is mortal," which he notes is an empirical proposition derived through the understanding without any direct reliance on reason (A322 / B378). Accordingly, reason's function is not to establish the truth of this proposition but to render it comprehensible by showing it to be a consequence of its condition (being human) taken in its full extent. As we have seen, it accomplishes this by finding a concept ("human"), which provides the condition under which the predicate ("mortality") can be affirmed of the subject (Caius). As we have also seen, however, in order to ground this predication it is necessary to take the condition in its whole domain, which is the function of the major premise ("All humans are mortal"). Moreover, since humans are judged mortal in virtue of being animals and animals in virtue of being living beings, and so on, it is obvious that this clarificatory process can be extended prosyllogistically.

This gives us the conception of a condition (being human, animal, or living being) taken in its totality. Here totality is just universality as a logical function (*universalitas*). Appealing to the results of the Metaphysical Deduction, however, Kant next reminds us that "[i]n the synthesis of intuition this [universality as a logical function] corresponds to **allness** (*universitas*), or the **totality** of conditions" (A322 / B379). Moreover, from this he concludes, without further argument, that "the transcendental concept of reason is none other than that of the totality of conditions to a given conditioned thing" (A322 / B379). Finally, assuming the reciprocity of the **totality of conditions** and the unconditioned, Kant states that "a pure concept of reason can in general be explained through the concept of the unconditioned, insofar as it contains a ground of synthesis for what is conditioned" (A322 / B379).

Setting aside for the moment the equation of reason's idea of the totality of conditions for a given conditioned with the unconditioned, what is bewildering



ing about this reasoning are the moves: first, from the concept of a condition taken in its totality (universality as logical function) to the concept of a totality of conditions (the category of allness or totality), and, second, from this to the totality of conditions for a given conditioned (the transcendental concept of reason). Since the plausibility of Kant's whole account hinges on the viability of these moves, they obviously require further consideration.

The initial move, from a condition taken in its totality to a totality of conditions, is clearly intended to reflect the move from logical function to category. At the level of general logic, where the former is operative, the concern is solely with the relation between the extensions of the constituent concepts. Accordingly, the extension of the concept is regarded as given in its totality. Or, equivalently, the condition is taken in its totality, as it must if the inference is to be valid. By contrast, the category, operative at the level of transcendental logic or the synthesis of intuition, is concerned with the determination of an extension in the first place. In the present case, this amounts to the complete collection of the *x*'s (Caius, Socrates, Mary, and so forth) falling under the concept "human." Thus, it involves the thought of the synthesis of the totality of the conditions for the predication expressed in the judgment.<sup>16</sup> 

The second move, from category to idea of reason, is more puzzling, particularly since Kant appears to present it as a direct consequence of the preceding characterization of the category of allness or totality. Indeed, in doing so Kant creates the impression that he is simply conflating two quite distinct senses of a "totality of conditions": the complete collection of the items falling under the extension of a concept, and the complete set of premises or explanatory grounds for a conclusion or matter of fact considered as conditioned.

On a more charitable reading, however, Kant can be taken as indicating the *distinction* between the category of allness or totality and a transcendental concept of reason, the very conception of which is here being explicated. Moreover, such a distinction is certainly needed, since he connects the conception of "totality" with both understanding and reason. So construed, Kant's point is that whereas "totality" as category is equivalent to "allness," in the sense of all the individuals falling under the extension of a concept, "totality" as thought by reason refers to the completeness of the set of conditions presupposed by something taken as conditioned. It follows from this that the former requires nothing more than an empirical, inductively based universality, while the latter includes all *conceivable* conditions.<sup>17</sup>

Kant underscores this crucial point by means of a reflection on the term *absolute*. Appealing to an ambiguity in its usage, according to which it can refer ei-

ther to what holds of a thing considered in itself or internally, as opposed to its being considered in relation to other things (for example, what is absolutely or inherently possible) or to what holds of a thing in all respects or in all relations, he proposes to use it in the latter sense (A324–25 / B381–82). Consequently, an absolute totality of conditions is one that constitutes a totality in all respects, which means that nothing can be added to (or subtracted from) it. Clearly, it is the totality of conditions in this sense that is of concern to reason and equivalent to the unconditioned.

Since this totality of conditions has reference to something conditioned, its thought necessarily makes use of the relational categories and consists in the extension of the relation between the conditioned and its conditions expressed in these categories to the unconditioned. In other words, the putative real or transcendental function of reason just is to extend this relation between the conditioned and its conditions, which is thought by the understanding in a piecemeal fashion, to the ideal goal of the totality of these conditions. Kant makes this connection between concepts of reason and the relational categories fully explicit when he remarks,

There will be as many concepts of reason as there are species of relation represented by the understanding by means of the categories; and so we must seek an **unconditioned**, **first**, for the **categorical** synthesis in a **subject**, **second** for the **hypothetical** synthesis of the members of a **series**, and **third** for the **disjunctive** synthesis of the parts in a **system**. [A323 / B379]<sup>18</sup>

Although this conception of ideas of reason may help to demystify Kant's position, it also suggests two closely related objections that have been raised in the literature. One concerns the apparent relativization or trivialization of the distinction between the understanding and reason as two forms of conceptualization. Rather than indicating that reason has a distinct function from the understanding, the above account suggests that the difference between them is merely a matter of degree: the understanding being assigned the responsibility for relatively low and reason for higher levels of generalization. But if this is Kant's view, then the distinction between understanding and reason is of questionable philosophical significance, unable to bear the heavy burden that he places upon it.<sup>19</sup>

Because it never leaves the level of understanding, which involves a hierarchy of rules but no genuine principles, such an objection completely misses the force of the notion of an *absolute* totality of conditions. Moreover, Kant emphasizes this point, insisting that reason's conception of the unconditioned is



completely alien to the understanding, since the latter, in virtue of its piecemeal procedure, which is itself a consequence of the necessity of underlaying its concepts with intuitions, can never regard *any* condition as absolute or unconditioned (A326 / B383). Indeed, we shall see that in following the demand of reason to think the unconditioned, the understanding is inevitably led to violate the conditions of its own legitimate employment.

The second line of objection concerns the connection between this genealogy of ideas of reason as products of the combination of the relational categories with reason's concept of the unconditioned and their supposed derivation from the forms of syllogistic inference. Kant's initial suggestion was that these forms provide the clue to the discovery of the ideas of reason, just as the forms of judgment do for the pure concepts of the understanding. But it now appears that Kant's true account of the origin of these ideas is quite independent of the structure of syllogistic reasoning. Moreover, though this has seemed to many a more promising approach, it raises questions about the consistency of Kant's position. As Kemp Smith puts it, "The deduction [of the ideas from the three forms of syllogism] is . . . wholly artificial, and masks Kant's real method of obtaining the Ideas, namely, through combination of the unique concept of the unconditioned with the three categories of relation."<sup>20</sup>

In order to understand the connection between these accounts, it is essential to keep in mind that Kant's initial suggestion was not that the syllogistic forms of themselves somehow yield ideas of reason but merely that they may do so, "if applied to the synthetic unity of intuition under the categories" (A321 / B378). Accordingly, we must determine what is meant by such a synthetic unity and how the syllogistic forms are supposedly applied to it.

To begin with, since it is the relational categories that are involved in the generation of ideas of reason, the synthetic unity in question must be governed by these. In other words, it consists in the thought of the relation of something conditioned to its condition, which is expressed in judgments having categorical, hypothetical, and disjunctive forms. But if they are to involve the thought of the unconditioned, these judgments must express a rule of the understanding in thinking these relations, that is, a pure use of a category. This because what makes a category extendable to the unconditioned is precisely its purity, that is, its lack of any reference to the spatiotemporal conditions of human cognition. Moreover, since the division of syllogisms into categorical, hypothetical, and disjunctive is itself based on the nature of the relation expressed in their major premises, the application in question must consist in the formation of syllogisms having these rules of the understanding as their major premises and



the assertion of an object corresponding to the idea (the objective reality of the idea) in its conclusion.

If this is correct, then a transcendental idea may be characterized as the conception of something unconditioned obtained through an inferential process embodying one of the syllogistic forms, in which the major premise is a judgment expressing a rule of the understanding regarded as a principle in the absolute sense.<sup>21</sup> This last qualification is crucial, since Kant defines reason as the faculty of principles and understands by these principles judgments that serve as the source of synthetic *a priori* cognitions without any appeal to *conditions* of intuition.<sup>22</sup>

This gives us the concept of a transcendental idea and tells us where to look for them, namely, the above-mentioned syllogisms in which the rules for the pure use of the understanding serve as a major premise. But it does not tell us either why the soul, world, and God should be included amongst these ideas or, more important from Kant's point of view, why they exhaust the set of them. This is the task of section 3 of Kant's introductory chapter on the ideas, which, like its counterpart in the *Analytic*, may be characterized as their metaphysical deduction proper.<sup>23</sup>

At the heart of this deduction lies yet another notorious trichotomy, which serves to define the set of generic relations in which all representations can stand: either to the subject or to objects, with the latter taken in two senses, namely, as appearances or as "objects of thinking in general." Moreover, Kant concludes from this that "all the relations of representations of which we can make either a concept or an idea are of three sorts: 1) the relation to the subject, 2) to the manifold of the object in appearance, and 3) to all things in general" (A334/B391).

Without informing the reader of the fact, Kant here shifts from a concern with the forms of the relations involved in cognition to their content, that is, the types of relations capable of yielding either a concept or an idea (but especially the latter). Nevertheless, it is clear that such a shift is required, if he is to specify the ideas resulting from the various inferential forms. Moreover, assuming that Kant is concerned only with the relations in which representations stand insofar as they enter into cognition, the initial dichotomy is unproblematic, since every such representation must stand in a relation to both a thinking subject and an object. The former follows from its nature as a representation, and the latter from its role in cognition.

Although the further division of the latter into relations to appearances and to things in general seems suspicious, since it suggests that Kant is surrepti-

tiously injecting transcendental idealism into the story, it need not be taken in this way. Kant can also be read as merely distinguishing between two conceptions of the conditioned objects of human cognition: one in which the conditioned is explicitly taken as something spatiotemporal; and the other in which it is considered merely in accordance with its concept as a thing (its spatiotemporal properties being ignored or abstracted from, though not denied). Clearly, the significance of this distinction depends on Kant's ability to show that each of these types of conditioned involves a distinct conception of the unconditioned; and though he will argue for just this in the course of the Dialectic, it is not obvious at this point. Consequently, like much else in this section, it is best taken as a promissory note.

For present purposes, however, the salient point is that this trichotomy gives Kant the basis for both the specification and systematic division of the transcendental ideas. Operating on the already established premise that these ideas are concerned with the unconditioned synthetic unity (absolute totality) of conditions for something conditioned, Kant suggests that they fall into three classes, "of which the **first** contains the absolute (unconditioned) **unity** of the **thinking subject**, the **second** the absolute **unity** of series of **conditions of appearance**, the **third** the absolute **unity** of the **condition of all objects of thought in general**" (A334 / B391). These forms of unconditioned synthetic unity are, in turn, identified with the soul, world, and God, respectively, which, as already noted, are the objects of the metaphysical disciplines of rational psychology, rational cosmology, and transcendental (rational) theology (A334–35 / B391–92).

Apart from the previously discussed connection between the concept of a transcendental idea and the forms of syllogistic inference, these identifications are the main reason for the extremely hostile reception accorded Kant's treatment of reason. The basic complaint, which appears in diverse forms, concerns Kant's apparent privileging of the ideas of the soul and God. Sometimes the emphasis is on what might be termed the "spiritualist tilt" this gives to his metaphysical scheme.<sup>24</sup> Elsewhere it is directed to what is perceived as its sheer arbitrariness: the underlying assumption being that, in slavish obedience to the requirements of his architectonic, Kant was led to inflate the received school philosophy of his time into the authentic dictates of pure reason.<sup>25</sup> Moreover, in light of this, it is frequently suggested as a more or less "friendly amendment" that Kant could have constructed a simpler and more coherent critique of metaphysics by framing the entire discussion in terms of an Antinomy.<sup>26</sup>

In response to the basic objection, here it must suffice to note that with his



tripartite division Kant is not attempting to provide an exhaustive inventory of all metaphysical positions, actual or possible. His concern is rather with a certain kind of metaphysical reasoning, namely, one that leads to the positing of transcendent entities. Since the ideas of the soul and God clearly fall within this category, there is nothing arbitrary in Kant's inclusion of them.

Moreover, Kant does not ignore the classical naturalistic metaphysical positions, for example, materialism and anti-theism. His procedure is rather to consider them under the rubric of cosmology, which seems entirely appropriate given the commitment of such positions to a denial of entities transcending the sensible world. Admittedly, this appears to support the critics' suggestion that Kant should have confined the Transcendental Dialectic to the Antinomy. Indeed, there is ample textual evidence, both from the *Nachlass* and the *Critique* itself, that Kant seriously entertained such a possibility.<sup>27</sup> Nevertheless, there are good reasons for Kant to adopt the more complex structure. First and foremost, it is required by his theory of reason. Although reason's conception of the unconditioned is illusory, it is not self-contradictory, as it would have to be if *every* attempt to think it yielded an antinomy. Rather, we shall see that an antinomy arises only when this attempt leads to a conflict between the laws of the understanding, which require that every condition be considered as itself conditioned, and reason's demand for the unconditioned. We shall also see that this occurs only in the case of the cosmological ideas.

Second, the ideas of the soul and God, as generally understood in the metaphysical tradition, cannot be dealt with adequately within the limited framework of rational cosmology. In the case of the former, this is largely because the simplicity attributed to the soul by rational psychology has a quite distinct genealogy from that affirmed in the Thesis of the Second Antinomy. Whereas the latter is arrived at by means of an analysis of the elements of something composite, the former is derived from a reflection on the nature of thinking. Similarly, the Fourth Antinomy, which deals with the concept of an absolutely necessary being, does not yield the concept of a being distinct from the spatiotemporal world as a whole.<sup>28</sup> But since the God of the tradition, in contrast to the God of Spinoza, is by definition extra-mundane, it is certainly reasonable to look outside the framework of cosmological thought (as understood by Kant) for the origin of the idea of such a being.

In conclusion, then, the extreme hostility with which Kant's introductory discussion of reason and its putative real use is frequently greeted is unwarranted. Although much of what Kant says here is obscure and promissory, it is neither arbitrary, uninteresting, nor incoherent. On the contrary, it contains a

deeply suggestive account of the rational grounds (as contrasted with psychological motives) underlying the endeavor to “think the whole,” or, more precisely, the various kinds of whole, that has traditionally been viewed as the project of metaphysics. Moreover, this account is closely linked with his conception of transcendental illusion, to which we now turn.

## II. TRANSCENDENTAL ILLUSION <sup>29</sup>

Kant’s preliminary discussion of transcendental illusion is complicated by the fact that it is intertwined with what purports to be a general account of error (supposedly including metaphysical error). This raises two problems, with which the first part of the present section will attempt to deal: (1) the meaning and adequacy of this initial account of error; and (2) its compatibility with the subsequent account of illusion. A second part will examine Kant’s controversial thesis that the illusion (though not the ensuing error) is both natural and unavoidable.

### A. Error and Illusion

Kant begins by noting that truth and error, as well as the illusion that leads to the latter, are to be found only in judgment, that is, in the relation between an intuited object of judgment and the understanding (A293 / B249–50). At first glance, this seems a straightforward appeal to the correspondence theory of truth. Truth arises when a judgment (the product of the understanding) agrees with its object (as given in sensible intuition), and error arises when it does not. Nevertheless, it is evident that Kant has a more complex picture in mind, since he locates the ground of this lack of agreement in neither the understanding itself nor in the senses but in “the unnoticed influence of sensibility on understanding, through which it happens that the subjective grounds of the judgment join with the objective ones, and make the latter deviate from their destination [*Bestimmung*]” (A294 / B350–51).

It is clear why Kant does not attribute error to either the understanding or the senses taken singly. It would be incoherent to attribute error to the understanding as such, since it contains the norms of correctness; and it cannot be attributed to the senses, since they do not judge at all. Because these are the only cognitive faculties involved in judgment, it follows that a general account of judgmental error (assuming such an account to be possible) must involve some sort of misfit between them, which is the point that Kant emphasizes in his characterization of an erroneous judgment of the understanding “as a diagonal

between two forces that determine the judgment in two different directions" (A295 / B351). What is not clear, however, is why such error must always be attributed to "the unnoticed influence of sensibility," and why this influence should be seen as a conflation of subjective and objective grounds of a judgment, which leads the latter to deviate from their "true destination."

Kant eliminates one potential source of puzzlement fairly easily. The obvious question raised by this characterization of error is its compatibility with the discursivity thesis, which entails that human cognition has a sensory component. Consequently, it might seem that the "influence of sensibility" would be more properly described as a condition of truth than as the source of error. Kant is keenly aware of the problem and addresses it in a note attached to the characterization, where he remarks:

Sensibility, subordinated to understanding, as the object to which the latter applies its function, is the source of real cognitions. But this same sensibility, insofar as it influences the action of the understanding and determines it to judgments, is the ground of error. [A294 / B351, note]

By distinguishing between the role of sensibility as subordinate to the understanding and as influencing the latter's actions, Kant is able to distinguish between its proper function and its role in generating error. Nevertheless, this hardly resolves all of the problems. One remaining difficulty concerns Kant's use of a physical model, involving conflicting "forces," each exerting an "influence" on judgment, to express the epistemic relation of competing grounds of belief.<sup>30</sup> Given Kant's fundamental distinction between the spontaneity of the understanding and the receptivity of the senses, this seems like a highly inappropriate metaphor for Kant to choose to illustrate his theory of error. Another, perhaps more severe problem concerns its plausibility as a general thesis about error. On even the most charitable reading, it obviously fails to account for things like errors of calculation (which presumably Kant would not view as judgmental error), but it is also not clear that it is adequate for understanding the kinds of error that are of interest to Kant in the *Critique*.

The latter are the errors allegedly committed by metaphysicians; but before turning to that topic we must note that, making allowance for its highly metaphorical nature, this account is applicable to more mundane types of error. A case in point, which is immediately suggested by the model, is error in ordinary perceptual judgment. For example, if I judge that a stick is bent because it appears as such when perceived in water, then my judgment is clearly being misled by the sensible appearance. Here the influence of sensibility is "unnoticed"

in the sense that it is not factored into the judgment. If it were I would recognize that the stick only *seems* to be bent because of the effect of water on light but that it really is not.

More interestingly, the model is also applicable to at least a certain kind of metaphysical error, namely, the sort treated in the *Dissertation* under the rubric of “subreptic axioms,” which involve a conflation of the conditions of sensible cognition (of things as they appear) with conditions of things as thought through the pure intellect (as they really are).<sup>31</sup> Here sensibility supposedly exerts a baneful influence on the understanding by leading it to take merely “subjective” conditions (sensible ones) as if they were “objective” conditions applicable to things as they are in themselves. In spite of Kant’s rejection of the assumption of the possibility of a purely intellectual cognition of things as they “really are,” this conception of metaphysical error survives in the *Critique*. That is why Kant assigns to the noumenon the role of a limiting concept, whose function is to “limit the pretension of sensibility” (A255 / B311). As we have seen, it remains important for Kant to preserve the critical conception of things as they are in themselves from any “contamination” by sensible, that is, spatiotemporal, predicates, even while denying that we can cognize things so considered.

Even granting this, however, it remains difficult to see how Kant could maintain that *all* metaphysical errors fit this mold. Indeed, in the *Amphiboly*, he appears to diagnose the fundamental errors of Leibnizian ontology as resulting from the failure to recognize the essential contribution to human cognition of sensibility and its *a priori* forms. How, then, can Kant in the very next chapter claim that metaphysical error is always to be viewed as the product of the “unnoticed influence of sensibility on understanding”? Or, alternatively, how can he claim that it always involves subjective grounds joining forces with objective ones leading the latter to “deviate from their destination”? From the Kantian point of view, the problem with the Leibnizian ontology would seem to be more accurately described as its denial of any role for “subjective” (sensible) conditions. But if sensibility has no positive epistemic function, how can it be said to join forces with “objective” (intellectual) conditions, leading the latter to deviate from their true destination?

Although the metaphor must be stretched somewhat, it does seem possible to see how Kant might characterize the errors of Leibnizian ontology in this way. In order to do so, however, it is necessary to view Kant’s account of error against the backdrop of the general critique of ontology presented in the *Analytic*. Simply put, the main point of this critique is that ontological claims rest on a transcendental employment of the categories, that is, their illicit employ-

ment (independently of their schemata) with respect to things in general, as contrasted with their legitimate empirical employment, which is limited to objects insofar as they are given in accord with the conditions of human sensibility. As we have seen, even though the categories are characterized as concepts "of an object in general," they yield cognition of objects only under the restricting conditions supplied by sensibility (the transcendental schemata). Accordingly, not to notice the influence of sensibility is to ignore either its subjective nature or its positive epistemic function with respect to the understanding. Moreover, it is precisely the recognition of these, which, according to Kant, requires that "the proud name of an ontology, which presumes to offer synthetic *a priori* cognitions of things in general . . . must give way to the modest one of a mere analytic of the pure understanding" (A247 / B303).

This suggests that there are two quite different, though complementary, ways in which the failure to notice the influence of sensibility on the understanding leads to metaphysical error, both of which involve an illicit extension of the categories to things in general, which may be corrected by a healthy dose of transcendental reflection. One is the way described above, whereby subreptic axioms direct the understanding to extend its principles to objects in general (thereby making a transcendental use of the categories), under the misguided assumption that sensible conditions are themselves ontological conditions of things in general rather than merely epistemic conditions of things as objects of possible experience. In the Amphiboly Kant attaches this form of error to the Lockean position, which he accuses of having "sensitized the concepts of the understanding" (A271 / B327). Clearly, however, this charge is intended to have a much broader scope than the views of the historical Locke. As already noted, Kant's recipe for avoiding this kind of error is to introduce the noumenon in order to "limit the pretension of sensibility."

The second way is that of Leibniz, who Kant contends "constructed an intellectual system of the world, or rather believed himself able to cognize the inner constitution of things by comparing all objects only with the understanding and the abstract formal concepts of its thinking" (A270 / B326). Although here the problem lies in a failure to notice the positive role of sensibility as the source of the *realizing* conditions of the understanding, the end result is the same, namely, the illicit extension of the categories to things in general. Moreover, this ontology may also be described as a case of subjective sources (sensibility) joining forces with objective ones (the understanding), thereby leading the latter to deviate from their true destiny. According to the discursivity thesis the true destiny of the understanding is to link up with sensible intuition,

thereby producing cognition. But by treating the sensibility-understanding distinction as one of degree rather than kind, Leibniz effectively conflated subjective and objective sources of cognition, which, in turn, led the understanding to deviate from its true destiny.

In sum, then, it can be argued that both species of metaphysician, albeit for diametrically opposed reasons, fail to notice the influence of sensibility on the understanding, and in both cases this leads to the illicit extension of the categories to things in general. The difference consists merely in the fact that in the first case it is based on the erroneous assumption that the conditions of sensibility are also conditions of things in general (which makes them "objective" in the Dissertation's sense), whereas in the second case it is based on the equally erroneous (though opposite) assumption that sensibility is not the source of any genuine conditions of cognition. Moreover, as should be clear by now, in both cases this failure is a direct function of their transcendently realistic assumptions, and in both cases the remedy lies in transcendental reflection.

If this is a correct interpretation of Kant's brief account of error (admittedly a highly controversial matter), then it is possible both to understand why he describes error in what initially seems to be such a misleading manner and to reconcile this account with the general critique of ontology sketched in the final portion of the *Analytic*.<sup>32</sup> At the same time, however, it also seems to exacerbate rather than resolve our second problem, which concerns the connection between metaphysical error, so construed, and the doctrine of transcendental illusion, which Kant introduces immediately following the account of error.

The question before us is, What role is illusion supposed to play in the generation of metaphysical error or, equivalently, why does Kant deem it necessary to introduce this additional complication, when he has presumably given an adequate account of metaphysical error in terms of principles taken from the *Analytic*? Indeed, since error thus analyzed (in both its subreptic and its Leibnizian forms) reduces to a misapplication of the categories and involves no reference to reason, one might wonder why a separate analysis of reason, that is, a transcendental dialectic, is necessary for the understanding and repudiation of metaphysical error?

The answer lies in Kant's underlying contention that there are two quite distinct kinds of illicit extension of the categories, each calling for a separate analysis and remedy. One, which has been described above, is their extension from objects of possible experience to things in general. The other is their extension from the conditioned objects of the understanding to the unconditioned objects of reason. Since the former pertains to what is traditionally known as on-

tology or general metaphysics (the science of being qua being), it does not involve the assumption of an additional sphere of objects or, therefore, any reference to reason. Consequently, it can be avoided by transcendental reflection, which means that the resources of the *Analytic* are sufficient both to expose and to avert it. This is not true, however, of the second form, which pertains to the traditional disciplines of special metaphysics and their putative objects (God, soul, and world). Since this involves the assumption of a distinct set of non-sensible objects, the concepts of which (as pertaining to the unconditioned) supposedly have their seat in reason rather than the understanding, it requires a separate and more extensive treatment.

Kant's introductory discussion of transcendental illusion is intended to explain the need for this separate treatment and thus to justify the inclusion in the *Critique* of the lengthy Transcendental Dialectic. Kant introduces this fresh conception of a transcendental illusion by contrasting it with the familiar optical variety. The basic claim is that, whereas the latter can infect the empirical use of concepts, transcendental illusion "influences principles whose use is not ever meant for experience, since in that case we would have at least a touchstone of their correctness." Instead, "contrary to all the warnings of criticism, [it] carries us away beyond the empirical use of the categories, and holds out for us the semblance of extending the pure understanding" (A295 / B352). As we have seen, however, this extension is quite different from the one diagnosed in the *Analytic*.

Although Kant is not yet in a position to make this clear, the key point is that this illusion has its basis in the very nature of reason. Thus, the reference is to "principles whose use is not ever meant for experience," which can mean only principles of reason, as contrasted with those of the understanding. Kant clarifies the difference by contrasting two types of principle: immanent ones, whose application is wholly within the boundaries of possible experience, and transcendent ones, which purport to bypass these boundaries. And he further notes:

By the latter I do not understand the **transcendental** use or misuse of categories, which is a mere mistake of the faculty of judgment when it is not properly checked by criticism, and thus does not attend enough to the boundaries of the territory in which alone the pure understanding is allowed its play; rather, I mean principles that actually incite us to tear down all those boundary posts and to lay claim to a wholly new territory that recognizes no demarcation anywhere. Hence **transcendental** and **transcendent** are not the same. [A296 / B352]

This distinction between the transcendental and the transcendent is one of the most important, yet most frequently misunderstood, distinctions in the

*Critique*.<sup>33</sup> As the above passage indicates, it amounts to a distinction between two kinds of principle: those whose proper use is immanent but which, as generic rules for the discursive thought of objects, contain in themselves no restriction to objects of possible experience, and those which actually demand that the understanding ignore any such restriction and project its sphere of normativity beyond the boundaries of possible experience. The passage also indicates that the remedy for the misuse of the former is proper attention to the boundary lines drawn by the *Critique*, that is, transcendental reflection. Thus, by distinguishing sharply between these two kinds of principle, Kant is suggesting that the harm caused by the latter cannot be dealt with so easily, which is why a Transcendental Dialectic is necessary.

Kant develops this point in the final two paragraphs of the section, which focus on the source and power of the transcendental illusion that supposedly underlies the appeal of transcendent principles. Once again, his procedure is to contrast it with another familiar kind of illusion, this time the logical variety. The latter consists merely in fallacious inferences that appear valid, but whose semblance of validity is entirely removed by proper attention to the rules of logic. Kant denies that this is true of transcendental illusion, however, since it remains even when uncovered by transcendental critique. Moreover, Kant suggests, this is because

in our reason (considered subjectively as a human faculty of cognition) there lie fundamental rules and maxims for its use, which look entirely like objective principles, and through them it comes about that the subjective necessity of a certain connection of our concepts on behalf of the understanding is taken for an objective necessity, the determination of things in themselves. [A297 / B353]

As this passage indicates, transcendental illusion consists in certain subjective principles of reason appearing to be objective. Although we shall soon see in more detail what this means, our immediate order of business is to underscore the difference between the present account of illusion and the previous account of error. The point, however, is not, as is sometimes thought, that they are two incompatible accounts of the same thing; it is rather that they address two fundamentally different sets of issues and that the failure to keep this in mind makes the understanding of the Transcendental Dialectic impossible. This difference is twofold. First, as already noted, they are concerned with different spheres of object: things in general, on the one hand, and a special set of transcendent objects, on the other. Second, this illusory appearance of objectivity is not of itself erroneous, though it can easily lead to error. It is one



thing for a subjective principle to appear objective (just as the stick in water appears bent) and quite another for it to be (mis)taken as such. Moreover, whereas in the case of judgmental error it was a matter of “subjective grounds” of a judgment (sensibility) joining with “objective grounds” (understanding), thereby inducing the latter to deviate from its true course, in the case of reason there is no deviation. On the contrary, in following these seemingly objective, though really merely subjective, rules and maxims reason is following its own natural course. Consequently, the illusion stems from the very nature of the principles involved, whereas judgmental error arises from a mistaken use of principles, which in their own sphere are perfectly legitimate. That is what makes exposing the “seat” of this illusion in reason and the errors it generates both so difficult and so important to the fate of metaphysics.

The appearance of objectivity generated by illusory principles also explains why Kant viewed transcendental illusion as both natural and unavoidable, even though judgmental error, including the variety stemming from such illusion, is neither. In developing this thought, Kant returns to the comparison with optical illusion. His point is that in both cases it is a matter of something seeming to be other than it is (Kant’s examples of the latter are the sea appearing higher at the horizon than at the shore, since it is seen through higher rays of light, and the rising moon appearing larger to an astronomer); and in both cases this seeming persists even after it is recognized as illusory. Moreover, as Kant points out, this unavoidability has important consequences for the treatment of metaphysical claims and the transcendent principles on which they are based. Specifically, it means that the most that can be accomplished in this domain is to uncover the underlying illusion and provide protection against being deceived by it (A297 / B354).

## **B. The Naturalness and Unavoidability of Transcendental Illusion**

In order to evaluate this bold and novel thesis concerning a natural and unavoidable illusion connected with rules and maxims of the use of reason, it is first necessary to determine the nature of the latter. Fortunately, in light of our earlier analysis of reason, this is easy to do. Clearly, the maxim in question can be none other than the one Kant designates for the logical use of reason, namely, “find the unconditioned for the conditioned cognitions of the understanding, with which its unity will be completed” (A307 / B364). Since this maxim expresses a necessary demand of reason, there can be nothing inherently illusory in it. Thus, the problem must lie in its inseparability from its compan-

ion: “[W]hen the conditioned is given, then so is the whole series of conditions subordinated one to the other, which is itself also given (i.e., contained in the object and its connection)” (A307–08 / B364).

Following Grier, these will henceforth be referred to as  $P_1$  and  $P_2$ , respectively; and, likewise following her, it will be argued that the key to understanding Kant’s doctrine of transcendental illusion lies in determining the relation between them.<sup>34</sup> Now, according to Kant,  $P_1$  becomes itself a principle of reason precisely by assuming  $P_2$ , which, as we have seen, he regards as obviously synthetic. This suggests that  $P_2$  serves as the application condition for  $P_1$ ; or, alternatively, that  $P_1$  and  $P_2$  stand in a relation that is analogous to that between a category and its schema and in that sense are reciprocals. The crucial difference, however, is that in the case of  $P_1$  the application condition cannot be met, since the absolute totality of conditions or, equivalently, the unconditioned posited by  $P_2$ , can never be given as an object.

We can see from this that the real use of  $P_1$ , that is, its use in connection with the thought of things and their conditions, must be inherently illusory, because such use is inseparable from an appeal to its application condition,  $P_2$ , and the latter is not realizable. Moreover, since, as a maxim for the logical use of reason,  $P_1$  is clearly a subjective principle (dictating how reason must proceed if it is to be consistent with its vocation), and since  $P_2$ , as a metaphysical assumption, purports to be objective, we can also see why this illusion consists in something subjective presenting itself as objective. It is not that  $P_1$  is simply conflated with  $P_2$ , since there remains a clear conceptual difference between them. It is rather that it cannot be acted upon without, at the same time, assuming  $P_2$ .

Nevertheless, it might still be doubted whether this illusion, which underlies the generation of the transcendental ideas considered in the first part of this chapter, is either natural or unavoidable. Indeed, this claim, like virtually every aspect of Kant’s account of reason, has been greeted with considerable skepticism. Once again, Walsh speaks for many when he remarks, “The illusion of which he speaks was perhaps ‘natural and inevitable’ to a thinker with Kant’s background in rationalist metaphysics, but would be less dangerous for, say, a scientifically-minded positivist.”<sup>35</sup>

Although Kant does not have much to say about the naturalness of this illusion, it is apparently due to the inseparability of  $P_2$  from  $P_1$ , which as the subjective principle of reason is itself clearly “natural.” At least it seems legitimate to characterize the illusion as such, if it is natural to assume that the conditions of the realization of a demand of reason hold. Consequently, we shall focus the discussion on the question of avoidability, which is the real bone of contention.

Given the preceding analysis of the illusion, there seem to be two main ways in which one might endeavor to avoid it. One is simply to refuse to recognize  $P_1$ , that is, to reject the whole project of seeking the unconditioned. The other is to attempt to drive a wedge between  $P_1$  and  $P_2$ , thereby obeying the "intellectual categorical imperative," while avoiding any of the metaphysical commitments this obedience is thought to involve. We shall briefly consider each in turn.

To begin with, it must be admitted that there is a clear sense in which this illusion is avoidable on the first strategy, though the significance of such an avoidance is highly questionable. One may do so simply by refusing to ask for conditions, which is analogous to avoiding an optical illusion by refusing to open one's eyes.<sup>36</sup> But since this is clearly not a particularly fruitful strategy for Walsh's scientifically minded positivist to adopt, it may be assumed that he would recommend the limitation of the inquiry to the search for particular conditions, that is, to the domain of the understanding, as Kant defines it, thereby eschewing any lofty but futile quest for the "unconditioned."

Although this may seem to be a plausible strategy, particularly for critics (like Walsh) who dismiss Kant's introduction of reason as a distinct faculty, it is easily shown to be deeply flawed. The problem is that this illusive unconditioned is equivalent to the totality of conditions for a given conditioned. Accordingly, what  $P_1$  requires may also be expressed in such seemingly less objectionable maxims as: never stop seeking conditions until one gets them all; never rest satisfied with an explanation that leaves something unexplained, and so on. Moreover, as scientifically minded, it is doubtful that Walsh's positivist could simply reject such injunctions. Employing Kantian terminology, it seems rather that the positivist would insist on the merely "regulative" status of these demands, thereby insulating the pursuit of scientific truth from any contamination by metaphysics.

At this point, however, the first of the aforementioned strategies for avoiding this illusion has mutated into the second. For what is now being claimed is precisely the separability of  $P_1$  and  $P_2$ . And, in support of this separability, it will presumably be argued that the search for conditions for a given conditioned, which characterizes the proper activity of the understanding with respect to experience, need not involve any illusory assumption about the "givenness" of the complete set of these conditions.

Once again, this seems to be a plausible response, perhaps even one that is genuinely Kantian. Nevertheless, once again, things are not that simple. The assumption underlying this response appears to be that a guarantee of success

in finding all the conditions is not a condition of the reasonableness of the search for them. The search might be partially successful, yielding some but not all of the sought-for conditions. This is perfectly correct, but beside the point. The question is not whether one must be assured of finding all the conditions for a given conditioned; it is rather whether it need be assumed that they are there to be found. But it is not at all clear that the latter assumption is dispensable. Indeed, precisely because the search is for *conditions*, it seems that the assumption (though not the search) cannot be abandoned without denying  $P_1$ . After all, a “condition hunt” is not like, say, the hunt for a hidden treasure, which one might reasonably pursue, while acknowledging that it may not exist. Since the conditioned is related analytically to *some* condition and that to its condition, and so forth, one cannot coherently endorse the possibility of there being something conditioned that lacks its sufficient conditions.

It seems, then, that transcendental illusion is not avoidable, and certainly not by the simple expedient of assuming an anti-metaphysical stance. In fact, we shall find that the positivist or putative anti-metaphysician, whom Kant regards as a dogmatic empiricist and identifies with the antithesis position in the various Antinomies, is really a metaphysician of a naturalistic sort and, therefore, does not escape falling prey to this illusion. We shall further see in the final chapter that the regulative use of reason, which the positivist is presumably willing to endorse, is itself based on this illusion.

What this shows is the importance of distinguishing between avoiding the illusion and avoiding its deceptiveness. As the analogy with optical illusion indicates, Kant’s thesis is that, though the former is unavoidable, the latter is not. Moreover, we shall see that the ultimate ground of this deceptiveness is the transcendental realism to which metaphysicians of all stripes (including the naturalistic sort) are implicitly committed. Consequently, in addition to exposing specific metaphysical errors, Kant’s therapeutic strategy in the Dialectic will involve the attempt to deprive this illusion of its power to deceive by removing the cancer of transcendental realism.<sup>37</sup> Although this strategy is most apparent in the Antinomies, where Kant offers an indirect proof of transcendental idealism by demonstrating the self-contradictory nature of the realism to which it is opposed, we shall find that it is operative in the Paralogisms and Ideal as well. And this, again, provides evidence of the inseparability of Kant’s overall project from transcendental idealism, properly understood.

## Chapter 12 The Paralogisms

The explicit concern of the Paralogisms is to provide a systematic critique of rational psychology. By the latter is understood the project of constructing a doctrine of the soul or self entirely on the basis of the meager resources of the *I think*, which Kant characterizes as its “sole text” (A343 / B401). The reason for this limitation stems from the fact that any further resources would necessarily involve empirical considerations and thus compromise the purity of the “science.” Although specifically associated with the philosophies of Wolff and Baumgarten, who distinguish sharply between rational and empirical psychology, its basis is clearly Cartesian.<sup>1</sup> Consequently, Kant’s critique of this procedure is part and parcel of his overall critique of what may be termed the “Cartesian project.”

Kant’s treatment of the Paralogisms, or at least the first three in the first edition, has been the subject of considerable interest, and it has been comparatively well received (particularly in contrast to other parts of the *Dialectic*) because of its affinities to contemporary debates in the philosophy of mind and philosophical psychology. Almost all of this contemporary discussion, however, even that which is historically

informed, has either completely ignored or marginalized the connections between Kant's critique of the claims of rational psychology and the underlying systematic concerns of the Dialectic, namely, transcendental illusion and the transcendental realism which, as previously noted, is the source of its deceptive power.<sup>2</sup>

This is hardly surprising, since what most contemporary Kant interpreters find interesting in the Paralogisms seems far removed from these systematic concerns, and it is only in connection with the Fourth Paralogism in A, frequently dismissed as an unfortunate venture into subjective idealism, that the topic of transcendental realism is explicitly discussed. Nevertheless, it will be argued that this approach is deeply mistaken, since the full force of Kant's critique of rational psychology can be appreciated only in light of these broader concerns. But since Kant rewrote the chapter in the second edition, it is necessary to consider this second account as well. Accordingly, the present chapter is divided into two main parts: the first deals with the first three Paralogisms in the A-edition, and the second with the revised account in the B-edition.

## I. THE PARALOGISMS IN THE FIRST EDITION

Since our main concern is to underscore the systematic issues that lie beneath the surface of Kant's argument, we shall focus mainly on the First Paralogism, which from the point of view of a theory of the self might seem of less intrinsic interest than the next two. It will also be argued, however, that the latter two, which deal, respectively, with the simplicity of the self (which is essential to the immateriality thesis) and its identity over time (personal identity), exhibit the same systematic structure.

### A. The First Paralogism

This Paralogism, which deals with the substantiality of the soul, consists in the following categorical syllogism:

That the representation of which is the **absolute subject** of our judgments, and hence cannot be used as the determination of another thing, is **substance**.

I, as a thinking being, am the **absolute subject** of all my possible judgments, and this representation of Myself cannot be used as the predicate of any other thing.

Thus I, as thinking being (soul), am **substance**. [A348]

According to Kant's official diagnosis, this inference commits the fallacy of a *sophisma figurae dictionis*, that is, a fallacy of equivocation manifested in an am-

biguous middle term. More specifically, the claim is that the major premise makes a merely transcendental use of the category of substance, whereas the minor premise and conclusion make an empirical use of the same category, subsuming the soul under the latter as the condition of the rule (A402). In other words, since "substance" is taken in different senses in the major and minor premises, the syllogism really consists of four terms, which obviously makes it invalid.

In his critical remarks on the First Paralogism, however, Kant focuses on the unfruitfulness rather than the invalidity of the inference, which has led to some confusion regarding the precise nature of Kant's complaint. The unfruitfulness is seen to stem from the fact that the major premise takes the category in its pure or transcendental sense as the concept of something in general, which, given the teachings of the *Analytic*, entails that it can admit of only a logical and not a real use. As Kant puts it, "Of any thing in general I can say that it is a substance, insofar as I distinguish it from mere predicates and determinations of things" (A349). Thus, one can perfectly well say this of the I, considered as thinking subject. But, as Kant points out, no real use can be made of this by the rational psychologist, since we are lacking the condition (the schema) under which alone the soul may be legitimately subsumed under the category. Consequently, no metaphysical conclusions concerning the immateriality, indestructibility, or persistence of the soul can be derived from it.

In spite of Kant's insistence to the contrary, this has led some commentators to deny that there is a formal fallacy in the syllogism as actually formulated and to look elsewhere for the rational psychologist's mistake. Thus, Bennett claims that the syllogism is perfectly respectable as it stands and that the error consists only in inflating its conclusion, that is, in taking it to demonstrate the immortality of the soul.<sup>3</sup> Similarly, Ameriks, though he takes issue with Bennett's claim that the argument really does establish that I am a substance, if this be construed merely as a formal truth, maintains the formal validity of the syllogism and suggests that a fallacy emerges only when the argument is extended to yield a putative conclusion regarding the soul's permanence. As Ameriks puts it, "Kant is defining the aims of rational psychology in terms of the prime practical-rational question of immortality, and so he simply does not attend to the claims of a modestly conceived rational psychology."<sup>4</sup>

Although such revisionary readings may be suggested by Kant's reflections on the First Paralogism and perhaps even by a superficial consideration of the syllogism (it does not immediately appear to be invalid), they cannot withstand closer scrutiny. To begin with the obvious, one would hardly expect the rational

psychologist's syllogisms to be transparently invalid, since they would then neither have the power to convince nor require a transcendental critique to expose. Thus, it must be assumed that some work will be required to locate the fallacy. Moreover, Kant specifies the fallacy of the First Paralogism in his critical discussion of it when he states that "it passes off [*ausgibt*] the constant logical subject of thinking as the cognition of a real subject of inherence . . ." (A350). Since the major premise merely provides what amounts to a nominal definition of substance (as pure category), it is clear that this "passing off" must occur in the minor premise.

Accordingly, the question is whether the minor premise may plausibly be taken to involve such an error, and the answer is clearly yes. Although this premise does not mention "the constant logical subject of thought," it effectively subsumes the concept of such a subject, in the guise of a representation of the I as "absolute subject of all my possible judgments," under the pure category of substance as characterized in the major premise. It should not be assumed, however, that the problem lies in a conflation of two senses of "absolute subject," namely, as ground of predication and subject matter of thought.<sup>5</sup> That would yield a fallacy of the type Kant claims, but one that is extremely uninteresting. Rather, the expression is taken in the first sense in both premises; but in the minor premise it is applied to the I of apperception in virtue of its function as logical subject, that is, as the ineliminable I think that "must be able to accompany all my representations" (as their constant logical subject). Consequently, the syllogism conflates the logical ineliminability of the I think with the real ineliminability (permanence) of the thinking subject as object.

In his concluding discussion of the Paralogisms, Kant describes this and the parallel fallacies committed by the rational psychologist in the other Paralogisms as "the subreption of hypostatized consciousness (*apperceptionis substantiate*)" (A402). As we saw in the preceding chapter, Kant understands by 'subreption' a metaphysical error stemming from a kind of conceptual confusion. More specifically, the confusion is of sensible and intellectual concepts and their respective domains (appearances and objects in general).<sup>6</sup> In the Dissertation, the subreptive axioms with which Kant was concerned involved the illicit extension of sensible concepts to things in general (the provenance of intellectual concepts). Consequently, Kant's account of subreption presupposes his sensibility-understanding distinction and, therefore, the discursivity thesis.

Given this, it seems relatively clear why Kant accuses the rational psychologist of a subreption involving the category of substance. The latter uses the category in its pure form in the major premise, while tacitly using its schematized



form in the minor. This accords with Kant's own account of the error as involving a "merely transcendental use of the category" in the major premise and an "empirical use" of the same category in the minor premise and conclusion (A402–03).<sup>7</sup>

Nevertheless, this still leaves us with three questions: (1) Why does Kant describe this subreption as of "hypostatized consciousness"? (2) How does the error, so described, relate to the underlying transcendental illusion? (3) What is the relation between this apparently substantive metaphysical error and the formal fallacy of the syllogism? The problem here is that such a misuse or "category mistake" appears to be a substantive metaphysical error rather than the purely formal one that Kant avers.

Since the last question is the easiest to deal with, we shall consider it first. The key point is that the metaphysical error underlying the syllogistic use of the premises does not affect the premises themselves. On the contrary, the minor premise, which alone comes into consideration at this juncture, is perfectly acceptable in its own right. The I, as thinking being, *is* the "absolute subject" of all its judgments and, considered as such, cannot be conceived predicatively. This is a consequence of Kant's account of apperception. Moreover, given the nature of the pure category, it likewise follows that the I, as thinking being, must *be conceived* as substance in this sense.<sup>8</sup> In order to generate the conclusion that this I *really is* a substance, however, the minor premise must involve an illicit substitution of the schematized for the pure use of the category, which means that the inference commits precisely the fallacy of which Kant accuses it.<sup>9</sup>

What makes the first two questions more difficult is that they concern the complex relationship of hypostatization to both subreption and illusion. If we identify the rational psychologist's metaphysical error solely with the subreption, then it appears that we must equate the hypostatization underlying this subreption with the illusion. Moreover, Kant himself seems to endorse this by suggesting the identification of the subreption of hypostatized consciousness with "the illusion of taking the unity in the synthesis of thoughts for a perceived unity in the subject of these thoughts" (A402).<sup>10</sup> But if this is what Kant means, then he is in contradiction with a basic tenet of the Dialectic. Since hypostatization is itself a form of metaphysical error, it must be avoidable and, therefore, cannot be identified with an unavoidable illusion. Accordingly, it seems necessary to regard hypostatization as an intermediate stage between illusion and subreption, thereby further complicating Kant's already complex account of the pathology of metaphysical error.

Although Kant's treatment of the topic is far from clear, what he says in the

Ideal and elsewhere indicates that he does view hypostatization as a form of metaphysical error, indeed the most basic form, which is intimately connected with transcendental realism. Kant understands hypostatization in fairly traditional terms as taking what merely exists in thought to be a real object existing outside the thinking subject (A384).<sup>11</sup> In the Paralogisms, it is obviously the I of apperception that is hypostatized by the rational psychologist; hence Kant's parenthetical use of the Latin expression "*apperceptionis substantiate*." Moreover, if by subreption is here understood an illicit use of the categories with respect to the thinking subject, it seems reasonable to claim that it presupposes a hypostatization. After all, it is only if the I of apperception is taken as an entity that one has any reason to apply the category of substance to it. It likewise seems to be the case, however, that one cannot hypostatize without also subrepting, at least not if one wishes to claim anything about what is hypostatized. Consequently, we are led to conclude that in the domain of rational psychology or, more generally, special metaphysics, hypostatization and subreption are reciprocal errors, which are nonetheless conceptually distinct.<sup>12</sup>

This brings us, then, to the question of the relation between the hypostatization of the I and the illusion. As already noted, they must be distinguished, since Kant distinguishes between fallacy and illusion. But since they are also intimately related, this becomes difficult to do. Moreover, the problem is exacerbated by the fact that Kant does not refer explicitly to the doctrine of illusion when discussing the individual Paralogisms, which no doubt helps to account for the widespread view that it is not of essential significance to the critique of rational psychology. He does point to the connection, however, in both the introductory section of the chapter common to both editions and the concluding portion of the A-Paralogisms, which is explicitly devoted to connecting the analysis with the doctrine of illusion (A396–403).<sup>13</sup> Indeed, in the very first paragraph of the chapter, Kant suggests that a transcendental paralogism (in contrast to a mere logical paralogism or ordinary formally invalid inference) "has a transcendental ground for inferring falsely due to its form." And he concludes from this that "a fallacy of this kind will have its ground in the nature of human reason and will bring with it an unavoidable, although not insoluble, illusion" (A341 / B399).

This suggests that the key to understanding the connection between the formal fallacies of rational psychology and transcendental illusion lies in the identification of this transcendental ground supposedly underlying these fallacies. Moreover, this ground cannot be identified with the principle of apperception itself or, more simply, the I think. Although the latter is clearly transcendental

and supposedly constitutes the “sole text of rational psychology,” it is that about which this science is under illusion rather than itself the source of the illusion. In fact, there is nothing in the I think as such to generate an illusion, which is to say that the principle of the synthetic unity of apperception is not inherently dialectical.

Here, as elsewhere, illusion arises only when reason enters the story, that is, when the attempt is made to find the condition of thought that is itself unconditioned.<sup>14</sup> This, in turn, means that  $P_2$  must be the sought-for transcendental ground.<sup>15</sup> Thus, the task is to understand how this illusory principle, which affirms the “givenness” of the absolute totality of conditions for any conditioned object of the understanding, leads the rational psychologist to hypostatize the I of apperception, which, in turn, leads to the illicit reasoning concerning the nature of this I.

In considering this question, we must keep in mind that Kant describes the illusion resulting from the use of  $P_2$  as a type of subjective-objective conflation. Specifically, “the subjective necessity of a certain connection of our concepts on behalf of the understanding is taken for an objective necessity, the determination of things in themselves” (A297 / B353).<sup>16</sup> Although not yet hypostatization, this is very close to it, which accounts for the difficulty in keeping them apart. Perhaps the best way of expressing the difference is to note that they involve two distinct types of subjective-objective conflation. In the case of the illusion it is a matter of conflating principles ( $P_1$  and  $P_2$ ). What has merely subjective validity as a maxim of reason ( $P_1$ ) is viewed as if it also had objective validity ( $P_2$ ). By contrast, hypostatization concerns not principles but putative entities. To hypostatize is not simply to take a merely subjectively valid principle as objectively valid but also, on the basis of this assumption, to assert a real existence. The hypostatization is based on the illusion without being identical to it.

Moreover, the latter is precisely what Kant accuses the rational psychologist of doing. To think the I of apperception as the absolute subject of thought, and, therefore, as “substance in concept,” as simple, and so forth, is subjectively necessary in virtue of its epistemic function as analyzed in the Transcendental Deduction. In this respect the rational psychologist is on solid ground. The problem, however, is that, under the influence of  $P_2$ , the same psychologist inevitably conflates this merely subjective necessity with an objective one and on that basis posits a real entity, the soul, answering to this description. As Kant notes in a *Reflexion* dated from 1778 or 1779, in the paralogistic illusion “the unity of apperception, which is subjective, is taken for the unity of the subject

as a thing" (R5533).<sup>17</sup> Or, as he puts it in the *Critique* itself, "[N]othing is more natural and seductive than the illusion of taking the unity in the synthesis of thoughts for a perceived unity in the subject of these thoughts" (A402).

What is strictly unavoidable, then, is to consider the unity of apperception as if it were the unity of a thing, because this is a condition of thinking of it as the unconditioned ground of thought. This is not the case with the fallacious hypostatization, however, since it involves an existential assertion that goes beyond this way of considering the unity of apperception. Accordingly, whereas the illusion is unavoidable, the hypostatization is not, and this holds also of the subreptic errors that follow from it.

Finally, this brings us to the thorny question of why the rational psychologist is deceived by this illusory objectivity, whereas the familiar kinds of optical illusion with which Kant compares it do not deceive anyone with the least knowledge of optics. The answer, as already indicated, lies in the transcendental realism tacitly assumed by the rational psychologist.<sup>18</sup> This realism, it will be recalled, consists in the conflation of appearances with things in themselves. Consequently, Kant's claim is that such conflation, which is itself natural (though not unavoidable) leaves one defenseless against the illusory objectivity of the unity of the thinking subject.

It is also not difficult to see why this is so. Simply put, transcendental realism does not have the conceptual tools for drawing the *kind of* subjective-objective distinction required to avoid being deceived by the illusion. This is not to say that it cannot draw *any* such distinction, since it is clear that transcendental realists distinguish all the time between what seems to be the case and what really is the case. Indeed, as I noted in the first part of this book, this underlies the standard misinterpretations of transcendental idealism. As we shall see in more detail in subsequent chapters, however, what it cannot do is distinguish between the merely subjective and objective *within the domain of the rationally necessary*. For such a realist, if a principle really seems to be objectively necessary in the sense of having the authorization of reason, then it must also be taken as fully objective and applicable to things as they are in themselves.

Transcendental idealism, however, does provide these tools, by allowing for (indeed insisting upon) the distinction between what is subjectively necessary for the *thought* of objects in general and what is objectively necessary for the *cognition* of objects of possible experience. This is not sufficient to remove the illusion, since  $P_2$  and its entailments continue to seem objective for the transcendental idealist, just as an optical illusion persists for someone with knowledge of optics. But they do make it possible to avoid being deceived by it,

which, as I have already noted, is Kant's aim. In the case of rational psychology, this occurs by revealing the hypostatization underlying the transcendental misapplication of the category of substance to the I of apperception.

Admittedly, the insidious role assigned to transcendental realism on this interpretation is not immediately apparent from the text. In fact, it is only in the Fourth Paralogism, which deals mainly with the problem of skeptical idealism rather than with the nature of the soul per se, that Kant appeals to it explicitly.<sup>19</sup> Nevertheless, it seems clear that Kant's ongoing polemic with transcendental realism underlies the critique of rational psychology, as it does the argument of the *Critique* as a whole.

## B. The Second and Third Paralogisms

Although from a strictly systematic point of view the Second and Third Paralogisms have little to add to this story, they are of considerable intrinsic interest and thus deserve separate consideration. Moreover, we shall see that Kant's treatment of them is perfectly in accord with the analysis offered above.

1. *The Second Paralogism.* This concerns the simplicity of the soul, which is closely related to its putative immateriality. It is expressed in the following dialectical syllogism:

That thing whose action can never be regarded as the concurrence of many acting things, is **simple**.

Now the soul, or the thinking I, is such a thing. Thus etc. [A351]

Significantly, Kant describes this as "the Achilles of all the dialectical inferences of the pure doctrine of the soul" (A351). Within the context of the analysis of transcendental illusion, this can refer only to its superior power to deceive, which may reflect the fact that Kant himself had at one time succumbed to it.<sup>20</sup> Indeed, he even now describes it as "an inference that seems to withstand even the sharpest testing and the greatest scruples of inquiry" (A351). Accordingly, Kant's goal is to show that this appearance of cogency is illusory.

Since the major premise merely provides a nominal definition of simplicity, the critique once again focuses almost entirely on the minor premise and, therefore, on the reasons, not provided in the inference, for attributing simplicity, so understood, to the thinking I. Although already advanced by Leibniz,<sup>21</sup> these reasons are closely related to Kant's account of apperception. In Kant's formulation, the argument for the minor premise consists in a *reductio* of the assumption that the thing which thinks is composite. This is deemed im-

possible on the grounds that the kind of unity required for thought (the unity of consciousness) cannot be conceived as the outcome of the collective action of distinct beings. As Kant puts it,

[B]ecause the representations that are divided among different beings (e.g., the individual words of a verse) never constitute a whole thought (a verse), the thought can never inhere in a composite as such. Thus it is possible only in one substance, which is not an aggregate of many, and hence it is absolutely simple. [A352]

This argument is clearly directed against the materialist, who, given the nature of matter as composed of parts externally related to one another, is supposedly committed to the thesis that thought can be understood as the collective product of such a composite. Thus, Kant locates the nerve of the *reductio* in the proposition that “many representations have to be contained in the absolute unity of the thinking subject in order to constitute one thought” (A353).

Although such an argument might not seem compelling from a contemporary point of view, since it relies on a rather narrow conception of scientific explanation,<sup>22</sup> its great affinity with Kant’s own account of the unity of apperception leads one to believe that it must have appeared so to him.<sup>23</sup> Moreover, since the major premise is non-problematic and the conclusion clearly follows from the two premises, the inference as a whole seems compelling. Consequently, the question becomes: What, from Kant’s standpoint, is wrong with the conclusion of the *reductio* and by extension with the minor premise of the Paralogism?

The matter is complicated by the fact that, rather than directly attacking the proposition at issue, Kant questions its grounds of proof.<sup>24</sup> In accordance with basic critical principles, he proposes three possible grounds and rejects all three. The rejection of two of these, namely, that this proposition can be cognized synthetically and *a priori* from concepts, and that it is derived from experience, are non-problematic. Kant also claims, however, that the proposition “‘A thought can be only the effect of the absolute unity of a thinking being’ cannot be treated as analytic” (A353). This is problematic because it appears to conflict directly with Kant’s insistence in the B-Deduction on the analyticity of the principle of the synthetic unity of apperception.

In order to avoid convicting Kant of a blatant inconsistency, it is necessary to distinguish between the proposition that here he denies to be analytic and the one that he affirms to be such in the B-Deduction.<sup>25</sup> Fortunately, this can be done, since, despite undeniable similarities, the proposition Kant attributes to

the rational psychologist differs from the principle of apperception in at least one crucial respect. The rational psychologist's claim, as Kant here presents it, is straightforwardly causal (it concerns the kind of entity of which thought might be "an effect"), and any such claim is obviously synthetic. Moreover, this is precisely the basis on which Kant criticizes the claim when, in denying the possibility of treating it as analytic, he remarks:

For the unity of a thought consisting of many representations is collective, and, as far as mere concepts are concerned, it can be related to the collective unity of the substances cooperating in it (as the movement of a body is the composite movement of all its parts) just as easily as to the absolute unity of the subject. [A353]

This diagnosis of the situation also enables us to understand what otherwise seems mysterious, namely, that Kant appeals to his doctrine of apperception in order to illustrate the *mistake* made by the rational psychologist. As in the First Paralogism, this mistake lies in the conflation of the unity of consciousness required as a logical condition of thought with the real or metaphysical unity (simplicity) of the thing that thinks. As Kant puts it in his conclusion to this part of the discussion of the Paralogism,

[T]he simplicity of the representation of a subject is not . . . a cognition of the simplicity of the subject itself, since its properties are entirely abstracted from it if it is designated merely through the expression "I," wholly empty of content (which I can apply to every thinking subject). [A355]

In view of the analysis of the First Paralogism, we can see that this conflation of the simplicity of the representation of the subject of thought with the simplicity of a real subject rests on the illusory hypostatization of this subject under the direction of  $P_2$ . We can also see that this leads naturally to an equivocal treatment of the minor premise and that the reason this equivocation is not noted lies in the rational psychologist's commitment to transcendental realism. Thus, if this argument is the "Achilles" of rational psychology, the latter is here, as elsewhere, its "Achilles' heel."<sup>26</sup>

2). *The Third Paralogism.* Here Kant is concerned for the first time in the Paralogisms with the diachronic dimension of consciousness, that is, with the self's consciousness of its identity through time or personal identity.<sup>27</sup> As such, it constitutes a parallel with the Second Paralogism, since simplicity may also be viewed as synchronic identity, that is, the identity of something with itself at a

time.<sup>28</sup> Once again, the argument under attack is given syllogistic form, with the major premise amounting to a nominal definition of personality and the minor premise, which subsumes the soul under this concept, bearing virtually all of the weight.<sup>29</sup> As Kant formulates the dialectical inference,

What is conscious of the numerical identity of its Self in different times, is to that extent a **person**.

Now the soul is etc.

Thus it is a person. [A361]

Since the basic analysis closely follows that of the Second Paralogism, we shall here focus on the features that distinguish it. Foremost among these is the contrast between the first- and third-person perspectives, which is introduced in the context of a distinction between the grounds of the ascription of numerical identity or sameness to objects of outer sense in space and to one's self as object of inner sense in time. In the former case, this ascription consists in an empirical judgment based on what is experienced to persist throughout a change of an object's determinations. In the latter case, however, there is no such empirical determination, even though the identity is ascribed to the self as object of inner sense. Rather, the claim is that the self cannot fail to find itself the same self throughout the time in which it is conscious of itself as object.

It follows from this that first-person judgments of identity have the indubitability attributed to them by Descartes and the rational psychologist, albeit on rather different grounds. For Kant, this is not to be understood as due to any unique access one has to an object (one's self); rather, it is merely because the thinking subject is necessarily for itself a single consciousness or I think. Thus, Kant suggests, "the personality of the soul must be regarded not as inferred but rather as a completely identical proposition of self-consciousness in time, and that is the cause of its being valid *a priori*" (A362).

Although it concerns the issue of numerical identity over time rather than the ownership of present mental states, Kant's point is reminiscent of Sidney Shoemaker's dictum, "A judgment of identity can be made only when a question of identity can be sensibly asked."<sup>30</sup> The rational psychologist assumes both that such a question can be posed with regard to the self and that a sufficient basis for an affirmative answer is provided from the first-person standpoint by the consciousness of an identical I think. Kant does not deny the possibility of raising the question of identity with regard to the self; but he does deny that it can be coherently either raised or answered from the exclusively



first-person perspective of the rational psychologist.<sup>31</sup> Accordingly, his basic tactic for exposing the rational psychologist's mistake is to contrast the first- and third-person standpoints on the identity of the self.

As already noted, the reason this identity is invariably met with from the first-person standpoint is that it is a formal condition of thinking, which, as such, indicates nothing about the nature of the thing that thinks. Thus, one might say that from this standpoint the question cannot be properly raised, since the ubiquitous and identical I is already presupposed. By contrast, from the third-person standpoint, wherein the question of personal identity is seen as similar in kind to the question of the identity of an object of outer sense, the question is perfectly in order. There, however, the I think is no longer available to answer it.

Kant attempts to illustrate this dual thesis by a pair of complementary thought experiments. In the first experiment (A363), one assumes a third-person standpoint with respect to oneself, that is, one regards oneself as an object of the outer intuition of an observer. Such an observer, Kant notes, will be perfectly willing to acknowledge that I necessarily find myself identical to myself throughout all time in which I am conscious of myself as an object of inner sense, because I accompany the thought of myself with an identical I think. Nevertheless, this has no bearing on the judgment of the observer, for whom I am merely an object of outer sense.

In the second experiment, Kant suggests the possibility of a scenario in which, analogously to the transference of motion from one body to another, consciousness is transferred from one thinking substance to another (A363–64). The point here is simply that under this scenario true, substantial identity (the kind insisted upon by the rational psychologist) would not be preserved, though the “logical identity of the I” would be, since it is “only a formal condition of my thoughts and their connection” (A363).

Accordingly, even though the situation is complicated by the contrast between first- and third-person perspectives on the self, the rational psychologist's error in the Third Paralogism parallels that of the first two. Specifically, it may be described as treating the question of the numerical identity of the self as a third-person type of question, that is, one concerning an object (albeit one given in inner sense), while attempting to answer it by appealing solely to first-person considerations concerning how the self necessarily conceives itself insofar as it refers its representations (this time of itself) to itself as the thinking subject. Moreover, this can again be seen as a “subreption of hypostatized con-

sciousness,” which is based on an underlying transcendental illusion that derives its power to deceive from the rational psychologist’s implicit commitment to transcendental realism.

## II. THE PARALOGISMS IN THE SECOND EDITION

As already noted, apart from the introductory section (A341–48 / B399–406), Kant completely rewrote the chapter on the Paralogisms for the second edition. Moreover, in the process, he provided a new characterization of the Fourth Paralogism, a much more concise account of the formal fallacy that the Paralogisms collectively embody, and an important discussion of the relation between apperception and the consciousness of the existence of the self, which completes the critique of the Cartesian conception of the *cogito* begun in the B-Deduction and the Refutation of Idealism.<sup>32</sup> Accordingly, the treatment of the revised version of the Paralogisms will again fall into two parts, the first dealing with the new analysis of the Paralogisms themselves and the second with the set of issues concerning apperception, existence, and the Cartesian *cogito*.

### A. The New Analysis

The key to Kant’s argument in the second edition lies in the claim that the Paralogisms as a whole are based upon the following invalid syllogism:

**What cannot be thought otherwise than as subject does not exist otherwise than as subject, and is therefore substance.**

**Now a thinking being, considered merely as such, cannot be thought otherwise than as subject.**

**Therefore it also exists only as such a thing, i.e., as substance. [B410–11]**

This calls to mind, but differs from, the syllogism associated with the First Paralogism in the first edition (A348). First, as is indicated by the inclusion of a reference to existence, the major premise no longer offers a merely nominal definition of substance. Instead, it states the condition under which some existing entity may be assigned substantival status, namely, that it can be thought only as subject. Second, the minor premise, rather than making an implicitly empirical (schematized) use of the category, claims simply that a thinking being, “considered merely as such,” meets this condition, from which it is inferred in the conclusion that it must exist in this manner. Thus, the new version of the Paralogisms brings to the fore a concern with the mode of existence of the thinking subject that was largely implicit in the original version.

Nevertheless, Kant continues to insist that the syllogism commits the fallacy of equivocation. Although in his analysis of this syllogism Kant suggests that the term used equivocally is 'thought', it seems more accurate to locate the equivocation in the whole expression: 'What cannot be thought otherwise than as subject'. As Kant points out, in the major premise this refers to an object or entity in general, and thus (supposedly) to something that can be given in intuition (B411). To say of such an entity that it cannot be thought otherwise than as subject is just to say that it is a substance. This is a synthetic judgment in which an object is subsumed under a category.

By contrast, in the minor premise it is simply the thinking subject, which cannot think itself otherwise than as a subject, insofar as it regards itself as the subject of thought. The key expression here is 'merely as such'. This is elliptical for 'merely as the subject of thought', which is the only way in which the rational psychologist claims to consider the self. Here the "thought" reduces to the tautology that the subject of thought must consider itself as the subject of thought. As such, it does not license the conclusion that this self-conscious subject of thought is an actual thinking substance. Consequently, in addition to the obvious hypostatization, it becomes clear in this new formulation that the rational psychologist is guilty of conflating a merely analytic proposition about how the subject of thought must conceive itself with a synthetic *a priori* proposition about the real nature of this subject.

By applying this schema to each of the inferences of rational psychology, Kant is able to diagnose the problem concisely, without attending pedantically in each case to the underlying syllogistic machinery. Thus, in the First Paralogism Kant notes that "in every judgment I am always the determining subject of that relation that constitutes judgment"; but he then proceeds to point out that the proposition "[T]he I that thinks can always be considered as **subject**, and as something that does not depend on thinking merely as a predicate . . . is an apodictic and even an **identical proposition**; but it does not signify that I as **object** am for myself a **self-subsisting being** or **substance**" (B407). This is because the latter claim, which is based on a hypostatization, is synthetic *a priori* and, therefore, must be sharply distinguished from the analytic claim to which the rational psychologist alone is entitled.

The Second Paralogism receives a similar concise treatment. Here Kant remarks, "That the I of apperception, consequently in every thought, is a **single thing** [*ein Singular*] that cannot be resolved into a plurality of subjects, and hence a logically simple subject, lies already in the concept of thinking, and is consequently an analytic proposition." But, he proceeds to point out, this does

not license the conclusion (drawn by the rational psychologist) “that the thinking I is a simple **substance**” (B407). Once again, the latter claim is synthetic and, as such, can be established only by an appeal to intuition, which in this case is unavailable.

The Third Paralogism, which affirms the numerical identity of the thinking subject, receives a virtually identical treatment. The basic point is that because of the illicit hypostatization the rational psychologist conflates an analytic proposition about the identity of the logical subject of thought with a synthetic proposition about the identity of a person over time, which, again, would require an appeal to intuition.

Since the Fourth Paralogism introduces a fresh topic, it requires a somewhat more extensive analysis. Its subject is the thesis that the soul, qua thinking being, can exist independently of the body. Thus, like its first-edition counterpart, it deals with an explicitly Cartesian doctrine. Unlike its counterpart, however, the issue is directly related to the central concerns of rational psychology.

The specific argument under attack infers the ontological distinctness of the soul or mind, qua thinking substance, from the body on the grounds that the subject can distinguish its own existence as thinking being from that of other things “outside” it, including its own body. Although Kant does not refer to Descartes here, this is readily recognizable as the well-known argument of the Second and Sixth Meditations. In the Second Meditation, Descartes contends that the mind is assured of its existence as a thing that thinks, even in the face of the assumption (based upon the Evil Genius) of the nonexistence of the body. Descartes recognizes that this does not of itself establish the distinctness of the mind from the body, but he does contend that this is established in the Sixth Meditation, where an appeal is made to the veracity of God. The argument therefore involves a combination of the claim that the mind can be clearly and distinctly conceived apart from the body (presumably established in the Second Meditation) with the principle “[E]verything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it.” On the basis of this, Descartes concludes:

Hence that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct, since they are capable of being separated by God. The question of what kind of power is required to bring about such a separation does not affect the judgement that the two things are distinct. Thus, simply by knowing that I exist and seeing at the same time that absolutely nothing else belongs to my nature or essence except that I am a thinking

thing, I can infer correctly that my essence consists solely in the fact that I am a thinking thing [or a substance whose whole essence or nature is to think].<sup>33</sup>

Although Kant's cryptic reformulation of this argument does not contain any reference to the familiar Cartesian notions of clear and distinct ideas and divine veracity, it does capture the essential move from separate conceivability to separate existence. Moreover, Kant characterizes the move in such a way as to make it clear that the error it involves is identical to the analytic-synthetic conflation committed in the other Paralogisms. Thus, he insists upon the analyticity of the claim that I can distinguish my existence as a thinking being from the existence of other things outside me, including my own body. The weight here falls on 'other things'. Since this refers to whatever is distinct from myself qua thinking being, it includes within its scope my own body. In this sense, then, it is an analytic truth that my body is "other" than, and thus distinct from, my mind. Nevertheless, the key point is that I cannot determine from this (as Descartes attempted to do) whether self-consciousness is possible "without things outside me through which representations are given to me, and thus whether I could exist merely as a thinking being (without being a human being)" (B409). Kant does not say so explicitly, but the latter is once again obviously a synthetic proposition, which requires an appeal to intuition.

At one point Kant suggests that this procedure of rational psychology is based on a mere misunderstanding [*ein blosser Missverstand*], namely, that "[t]he unity of consciousness, which grounds the categories, is here taken for an intuition of the subject as an object, and the category of substance is applied to it" (B421–22). Since a misunderstanding is clearly distinct from an illusion and presumably avoidable, it might be wondered what has happened to the thesis that the paralogistic inferences are grounded in an unavoidable transcendental illusion. Moreover, the problem is exacerbated by the fact that the misunderstanding Kant describes is clearly an instance of hypostatization cum subreption. Near the end of his discussion, however, Kant turns to this topic, making clear that he still considers the doctrine of illusion essential to the understanding of the Paralogisms. As he now puts it,

The dialectical illusion in rational psychology rests on the confusion of an idea of reason (of a pure intelligence) with the concept, in every way indeterminate, of a thinking being in general. I think of myself, in behalf of a possible experience, by abstracting from all actual experience, and from this conclude that I could become conscious of my existence even outside experience and its empirical conditions.

Consequently I confuse the possible **abstraction** from my empirically determined existence with the supposed consciousness of a **separate** possible existence of my thinking Self, and I believe that I cognize what is substantial in me as a transcendental subject, since I have in thought merely the unity of consciousness that grounds everything determinate as the mere form of cognition. [B426–27]

Here the illusion is said to consist in the identification of the indeterminate concept of a thinking being (a unified consciousness), which is arrived at by abstracting from all the content of thought, with the idea of a transcendent entity, here defined as a “pure intelligence.” This identification is illusory because reason has no such object or, better, has it “only in idea.” It is unavoidable because, under the influence of  $P_2$ , the indeterminate concept of a thinking being, which results from an abstraction from all the content of thought, presents itself as the concept of such a pure intelligence. Although the language differs from that of the first edition, the basic thought remains the same. In both cases, it is matter of confusing (because of  $P_2$ ) the subjectively necessary unity of thought with an objectively necessary condition of this unity.

Once again, the illusion concerns the epistemic status of a thought (assigning it objective rather than merely subjective validity). Consequently, it is distinct from the hypostatization, which, as an attribution of real existence, is itself the foundation of the subreptic errors committed by the rational psychologist. But since it is precisely because the thought is taken to have objective validity that the rational psychologist hypostatizes its putative object, the hypostatization is a direct consequence of the illusion. Indeed, it is the inevitable result of being taken in by it, though not of the illusion itself. As such, it is likewise equivalent to the “misunderstanding” described above, according to which the unity of consciousness is taken for an intuition of the subject as object, and the category of substance is applied to it.

Finally, even though Kant does not make it explicit, succumbing to this illusion may once again be seen as due to the transcendental realism of the rational psychologist. Moreover, this helps to explain Kant’s otherwise puzzling claim regarding the rational psychologist’s conflation of “the unity of consciousness, which grounds the categories,” with “an intuition of the subject as an object.” For Kant, of course, we have no intuition of the subject as object or, more precisely, the inner intuition we do have provides merely the contents of inner sense and not an abiding subject to which this content belongs. But since transcendental realism (because of its endemic failure to draw the sensibility-understanding distinction in the correct way) is incapable of distinguishing between the conditions of the *thought* of the subject and the conditions of its

*intuition* as an object, it unavoidably conflates these.<sup>34</sup> And the end result is a further conflation of an analytic proposition about the former with a synthetic one about the latter.

## B. Apperception and Existence

Given the above analysis of the B-Paralogisms, which emphasizes both the error of hypostatization and the systematic conflation of analytic propositions about the subject of thought considered as such affirmed in the minor premises and the synthetic (because existential) nature of the conclusions drawn by rational psychology, it might seem that Kant wished to separate his doctrine of apperception from any connection with claims about the existence of the self. In reality, however, Kant's position is much more nuanced, since, on the one hand, he sharply distinguishes his view from the Cartesian position, while, on the other, he includes an existential dimension in his account of apperception. In fact, already in the B-Deduction, when contrasting apperception and inner sense as modes of self-consciousness, Kant remarks that in the former "I am conscious of myself not **as** I appear to myself, nor as I am in myself, but only **that** I am" (B157). Moreover, in a note he adds:

The **I think** expresses the act of determining my existence. The existence is thereby already given, but the way in which I am to determine it, i.e., the manifold that I am to posit in myself as belonging to it, is not yet thereby given. For that self-intuition is required, which is grounded in an *a priori* given form, i.e., time, which is sensible and belongs to the receptivity of the determinable. [B157, note]

This theme is also carried over to the B-Paralogisms, where Kant prefaces his critique of the paralogistic inferences with the contrast between the mere thought or consciousness of self and genuine cognition. The former is connected with the determining and the latter with the determinable self (B407). By the former Kant understands the active, spontaneous, epistemic subject or, more simply, the self qua synthesizer, and by the latter the self that is cognized through inner intuition or introspection and whose existence is determined in time. As we have seen, Kant is committed to the possibility of the *consciousness* of the former self, that is, a consciousness of the act of thinking, by his account of the discursive nature of cognition, and to the denial of the possibility of a *cognition* of this self by the intuition requirement. But, in spite of this denial of any genuine cognition of the self through this mode of consciousness, Kant insists that it involves an assurance of existence ("that I am").

Nor are these isolated passages. Thus, in commenting on the organization of

his critique of the doctrines of rational psychology, Kant notes that if one follows the “analytic procedure” (which is that of Descartes in the *Meditations*), rational psychology begins with “the ‘I think’ given as a proposition that already includes existence in itself” and, therefore, “not from the concept of a thinking being in general but from an actuality” (B418). And, later, in the context of a discussion of materialism, Kant remarks that “apperception is something real, and its simplicity lies already in its possibility” (B419). Finally, in a lengthy and notoriously obscure footnote, which also contains the gist of his critique of Descartes’s *cogito* inference, he writes:

The “I think” is, as has already been said, an empirical proposition, and contains within itself the proposition “I exist.” But I cannot say “Everything that thinks, exists;” for then the property of thinking would make all beings possessing it into necessary beings. Hence my existence also cannot be regarded as inferred from the proposition “I think,” as Descartes held (for otherwise the major premise, “Everything that thinks, exists” would have to precede it), but rather it is identical with it. It expresses an indeterminate empirical intuition, i.e., a perception (hence it proves that sensation, which consequently belongs to sensibility, grounds this existential proposition), but it precedes the experience that is to determine the object of perception through the category in regard to time; and here existence is not yet a category, which is not related to an indeterminately given object, but rather to an object of which one has a concept, and about which one wants to know whether or not it is posited outside this concept. An indeterminate perception here signifies only something real, which was given, and indeed only to thinking in general, thus not as appearance, and not as a thing in itself (a noumenon), but rather as something that in fact exists and is indicated as an existing thing in the proposition “I think.” For it is to be noted that if I have called the proposition “I think” an empirical proposition, I would not say by this that the I in this proposition is an empirical representation; for it is rather purely intellectual, because it belongs to thinking in general. Only without any empirical representation, which provides the material for thinking, the act I think would not take place, and the empirical is only the condition of the application, or use, of the pure intellectual faculty. [B422, note]

For present purposes at least, the main puzzle raised by these texts (particularly the note) is to understand how this claim about the existence of the “apperceptive self” (if such an expression is permitted) is compatible with Kant’s attack on the supposedly illicit hypostatization of the self by the rational psychologist. Is not Kant himself here guilty of the same error of which he accuses the latter? Although Kant does not say so explicitly, the above-cited note may be read as containing his attempt to deal with this question within the context



of the juxtaposition of the critical view of the relation between the *cogito* and existence with that of Descartes. This attempt involves three distinct, though closely related, elements: (1) the claim that existence is already given in the “I think,” or equivalently, that it “contains within itself the proposition ‘I exist’”; (2) the claim that “I think” is an empirical proposition; and (3) the critique of the Cartesian inference, *cogito, ergo sum*. In what follows we shall consider each of these separately, though of necessity there will be some overlap.

1. *“I Think” and “I Exist.”* Although Kant’s account of the nature of the connection between these two propositions is obviously related to his critique of Descartes, the logical point he is trying to make is independent of this critique. In the first edition, he makes it with specific reference to Descartes by declaring that what is referred to as the Cartesian inference, *cogito, ergo sum*, is really a tautology, “since the cogito (*sum cogitans*) immediately asserts the reality” (A355). In the second-edition footnote cited above, he indicates that his point is simply that the propositions “I think,” “I am thinking,” and “I exist thinking” are all equivalent. Kant’s claim, then, seems to be that the existential assumption or presupposition is already built into the proposition “I think” and, therefore, cannot appropriately be regarded as an inference from it.

In an interesting discussion of Kant’s cryptic remarks concerning the *cogito*, Bernard Williams has suggested as a possible source for Kant’s view the dictum of Spinoza: “‘I think, therefore, I am’ is a proposition equivalent to ‘I am thinking.’”<sup>35</sup> Williams is certainly correct in his claim that this expresses Kant’s view. Nevertheless, given Kant’s extremely negative attitude toward Spinoza, this does not appear to be a likely source. Far closer to home is Leibniz, who makes the same point in the *New Essays* when he writes, “To say *I think therefore I am* is not really to prove existence from thought, since *to think* is already to say *I am*.”<sup>36</sup> Since Kant was a keen student of the *New Essays*, he was undoubtedly familiar with this remark of Leibniz’s. There is, however, really no need to assume any external source for Kant’s view of the *cogito*. The main point is simply that his claim about existence follows directly from the conception of apperception as a consciousness of the *activity* of thinking. To be aware of oneself as engaged in this activity is to be aware of oneself as existing, though it does not bring with it any further knowledge of one’s nature as a thinking being. Presumably, this is also why Kant would deny that it involves the hypostatization of a substantial self. Here the basic point is simply that such hypostatization rests upon a determinate conception of the self as a thinking being (which is what gets hypostatized) and, therefore, on the underlying illusion.

2. "*I Think*" as an Empirical Proposition. This claim is more difficult, particularly if one tries to connect it with the contention that the I that functions as the subject of the proposition is not itself empirical. Certainly, part of what Kant means here is that the proposition expresses a contingent rather than a necessary truth. In this respect he is again in agreement with Leibniz, who contends that it is a "proposition of fact, founded on immediate experience, and is not a necessary proposition whose necessity is seen in the immediate agreement of ideas."<sup>37</sup> The text, however, also indicates that part of what Kant means is that some given sensible representation must function as the occasion for the act of thought and, therefore, for the awareness of existence. Without something given to sensibility, that is, without sensation, there would be no *cogitatio* or, better, no *ergo sum cogitans*. In other words, the apprehension of some sensible content (as modification of inner sense) is a necessary condition of the awareness of existence that is presumed to be inseparable from the consciousness of thinking. Thus Kant maintains parenthetically that "sensation, which consequently belongs to sensibility, grounds this existential proposition," and, again, at the very end of the note, that "the empirical is only the condition of the application, or use, of the pure intellectual faculty." The relevant point here is not that the empirical is *only* the condition but rather that it *is* the condition, since it follows from this that apperception as an actual consciousness of thinking ("something real") always involves an empirical element.

Of equal importance, however, is Kant's insistence that the sensation that provides the occasion for apperception, and, therefore, for the apprehension of existence, is not itself an empirical representation of the subject. In fact, it is not a representation of the subject at all. Moreover, since in pure or transcendental apperception one explicitly abstracts from everything empirical, including sensation, and since there is no determinate non-empirical representation of the subject (no intellectual intuition), it follows that the thinking subject, whose existence is given or "contained" in the consciousness of thinking, can be characterized only as a "something in general =  $x$ ." This is clearly a non-empirical, "purely intellectual" representation: indeed, it is nothing more than the empty thought of a logical subject. Consequently, it does not follow from the fact that the thought of an existing subject of thought is non-empirical that we can have any non-empirical knowledge of the real nature of this subject. But the latter is precisely the assumption underlying the rational psychologist's illicit hypostatization of the self.

This also puts us in a position to interpret Kant's mysterious remark that the

notion of existence involved in the *cogito* cannot be equated with the category. Kant justifies this on the grounds that we are dealing with an “indeterminately given object,” and the category does not apply to such an object “but rather to an object of which one has a concept, and about which one wants to know whether or not it is posited outside this concept.” Behind this lies Kant’s account of the peculiar nature of the modal functions and categories, which, as we have seen, do not add anything to the content of the judgment.<sup>38</sup> Given this account, it follows that the category of existence is called into play only when we have a determinate concept and wish to determine whether there is an actual object answering to it. And since to hypostatize just is to apply the category of existence to a putative object of thought, it follows that where there is no such object there can be no hypostatization. But this is precisely what is lacking in the case of the I of the “I think.” Instead of a determinate concept of a thinking subject, we have only an “indeterminate perception” or, as Kant elsewhere puts it, a “bare consciousness” (A346 / B404) that is inseparable from, or “contained in,” the act of thought. Consequently, in connecting the latter with existence, Kant, unlike the rational psychologist, neither uses the category nor is guilty of a hypostatization.

3. *Cogito, ergo sum: The Kantian Critique.* Although Kant’s critique of the Cartesian *cogito* inference is embedded in his general critique of rational psychology, it is more convenient to treat it separately. The critique itself falls into two parts. The first is based upon Kant’s controversial interpretation of the inference as a syllogism. His claim is that, so construed, it generates the absurd conclusion that whatever thinks, exists necessarily. The second, which is closely connected with the general argument of the Paralogisms, is directed against the Cartesian project as it is manifest in the whole program of radical doubt. Seen through Kantian spectacles, this project can be described as the attempt to arrive at certainty about the existence of the self as *res cogitans* by simply reflecting upon what must be presupposed as a condition of thinking.

As is often pointed out, the basic problem with the first part of Kant’s critique lies in his syllogistic reading of the *cogito* inference, a reading that is widely rejected by Descartes’s interpreters.<sup>39</sup> But since Kant bases his critique of Descartes on this point entirely upon his own account of apperception, it is independent of his interpretation of the logical form of the *cogito* inference. His central claim is that the *cogito*, which survives the program of radical doubt, is nothing more than the empty “I think.” Moreover, the reason it survives is sim-

ply that it must be able to accompany all of my representations. Correlatively, our conception of this *cogito* must be characterized as the “bare representation” or “purely intellectual consciousness” of an abiding subject of thought.

Given this, Descartes’s mistake is clear: he identified this formal or transcendental I with a real self, that is, he hypostatizes it. Moreover, it is because of this hypostatization that Descartes erroneously believed that he had arrived by means of the *cogito* inference at a certainty regarding his own existence as a particular thinking substance (*res cogitans*). Finally, it is also clear that this error is rooted in the very transcendental illusion that underlies the entire project of rational psychology. As a dedicated transcendental realist, Descartes, like other rational psychologists, had no philosophical weapons to protect himself against this unavoidable illusion. Thus, he inevitably succumbed to it in his positive claims about the *cogito* as well as in his doubts about the reality of the external world (which for Kant are really two sides of the same coin). In the next chapter we shall see how, in the domain of rational cosmology, the same lethal combination of transcendental illusion and transcendental realism leads reason into an antinomial conflict, which can be resolved only by appeal to transcendental idealism.

## Chapter 13 The Antinomy of Pure Reason

It is virtually impossible to overestimate the importance of the Antinomy to Kant's critical project. Indeed, it is clear that Kant himself saw it as absolutely central. Thus, in a famous letter to Christian Garve from 1798, he writes that it was the Antinomy of Pure Reason that "first aroused me from my dogmatic slumber and drove me to the critique of reason itself in order to resolve the ostensible contradiction of reason with itself" (Br 12: 258; 552).

Lest it seem strange to find Kant here using virtually the same language to describe the discovery of the Antinomy as he had to characterize the effect of his "recollection of David Hume" some fifteen years earlier (Pro 4: 260; 57), it should be kept in mind that Kant also states in the *Prolegomena* that "[t]his product of pure reason in its transcendent use is its most remarkable phenomenon, and it works the most strongly of all to awaken philosophy from its dogmatic slumber, and to prompt it toward the difficult business of the critique of reason itself" (Pro 4: 338; 129). In fact, already in the first edition of the *Critique* Kant had written that the Antinomy "guards reason against the slumber of an imagined conviction, such as a merely one-sided illusion produces" (A407 / B434).

Because of its great systematic significance, the variety and complexity of the issues involved, as well as its intimate connection with transcendental idealism, Kant's account of the Antinomy requires a much fuller discussion than we gave to the Paralogisms. Nevertheless, anything approaching an adequate treatment of the four distinct cosmological conflicts or "antinomies" into which Kant divides his analysis of the Antinomy of Pure Reason would require several chapters, if not an entire book.<sup>1</sup> Accordingly, in order to keep the discussion within reasonable bounds, after an introductory account of the problematic and method of the Antinomy as a whole, the second and third parts of this chapter will focus, respectively, on the First and Third Antinomies. The fourth and final part will then deal with the crucial claims that transcendental idealism provides the key to the resolution of the Antinomy and that the analysis provides an indirect proof of this idealism, which complements the direct proof offered in the Aesthetic.

### I. THE ESSENTIAL PRELIMINARIES

The Antinomy deals with the species of illusion connected with the cosmological idea of the sensible world, understood as the sum total of appearances (spatiotemporal entities and events). As Kant points out at the very beginning of his discussion, the great significance of this illusion (as contrasted with those connected with the psychological and theological ideas) is that it is two-sided. In other words, reasoning based on this idea (and the various cosmological ideas through which it is expressed) yields two equally compelling, yet apparently contradictory, sets of conclusions. Since the equal success of each side in refuting the other precludes the possibility of a dogmatic solution to the conflict, Kant claims that the situation poses a dilemma consisting of a choice between a radical skepticism regarding reason and a stubborn dogmatism that simply ignores the equal cogency of the opposed assertions.<sup>2</sup> Moreover, Kant suggests, either option leads to "the euthanasia of pure reason" (A407 / B434).

Whereas Pierre Bayle, who had advanced similar sets of conflicting arguments a century earlier,<sup>3</sup> opted for the skeptical horn, Kant's grand strategy for avoiding the dilemma altogether is to adopt what he calls the "skeptical method," which consists in an examination of the presuppositions underlying the dispute (A424 / B451–52). Not surprisingly, this examination reveals that the conflict is based on a transcendental illusion shared by all parties, which combined with the pernicious influence of transcendental realism inevitably

leads reason into this morass. This, in turn, makes possible a “critical resolution” of the conflict by denying its underlying presupposition. Since it provides the basis for Kant’s entire account, this analysis of the nature and source of the conflict will be our initial concern. The discussion is divided into three parts. The first considers the nature of this two-sided illusion and the fallacy it engenders; the second discusses Kant’s attempt to show how the same illusion gives rise to the specific cosmological ideas; and the third analyzes his claim that the antinomies arising from these ideas fall into two classes, each of which requires a distinct manner of resolution.

### A. Illusion and Fallacy in Rational Cosmology

Kant claims that the Antinomy as a whole rests on this dialectical argument: “If the conditioned is given, then the whole series of all conditions for it is also given; now objects of the senses are given as conditioned; consequently, etc.” (A497 / B525). Since the major premise of this hypothetical syllogism is just  $P_2$ , Kant is here explicitly connecting the problematic of the Antinomy with the doctrine of transcendental illusion. Applying this dialectical syllogism to the idea of the sensible world yields antinomies rather than merely paralogisms because there turn out to be two incompatible ways of thinking such a totality, each of which accords with reason’s idea of the unconditioned: as containing a first, unconditioned condition (the thesis) and as containing an infinite number of conditions, only the totality of which is unconditioned (the antithesis). For reasons that will become clear shortly, Kant himself characterizes the former as the “dogmatism of pure reason” and the latter as “pure empiricism.” He further associates them, respectively, with Plato and Epicurus (A471–72 / B499–500).

The reason why such a totality can be thought in these two ways lies in the peculiar nature of the cosmological idea (the idea of the world). As the idea of the totality of *appearances*, it differs from the other transcendental ideas in that it does not seem to refer to a transcendent entity. In the *Critique* Kant puts the point by remarking, “The cosmological ideas alone have the peculiarity that they can presuppose their object, and the empirical synthesis required for its concept, as given” (A478–79 / B506–07). In the *Prolegomena* he notes, more expansively, that, insofar as a cosmological idea’s object is sensible it, unlike the soul, is immanent rather than transcendent and thus “up to this point . . . not yet an idea.” Nevertheless, it becomes an idea (in the technical sense) because it “expands the connection of the conditioned with its condition . . . so greatly

that experience can never match it, and therefore it is, with respect to this point, always an idea whose object can never be adequately given in any experience whatever" (Pro 4: 338; 129).

This in effect means that whereas the ideas of soul and God may be viewed as pseudo-rational, since they involve the illusion of referring to a non-sensible object of pure reason, the idea of the world may be viewed as pseudo-empirical, since it involves the illusion of referring to a high-order sensible object. As such, however, the idea of the world is governed by two conflicting norms, those of the understanding and of reason. Insofar as it purports to be *empirical*, it is governed by the former, which means that it is subject to the conditions of the possibility of experience. Insofar as it purports to be an *idea*, it is governed by the latter, which means that it involves the thought of the unconditioned. Consequently, it is precisely its subjection to these two conflicting norms that accounts for the two-sidedness of the illusion connected with this idea and the inferences based upon it. Each of the positions delineated above adheres consistently to one of these norms, eschewing the other, which is why Kant characterizes them as the "dogmatism of pure reason" (Plato) and "pure empiricism" (Epicurus).

This fundamental conflict is played out in each of the Antinomies. Kant claims that in each case the problem stems from the fact that the completeness or totality of conditions required by reason is "too large for the understanding," since it can never be given in a possible experience, whereas the relation to possible experience required by the understanding is "too small for reason," since it can never yield completion or totality (A422 / B450). At first glance, this characterization of the conflict might suggest that victory simply be assigned to the empiricistic position on the grounds that it represents the already legitimated demands of the understanding. We shall see, however, that this ignores the crucial fact that pure empiricism is as dogmatic in its own way as its intellectualistic opponent, since it too is under the spell of  $P_2$  and, therefore, assumes that it is referring to a given totality of conditions, which, as such, is unconditioned.<sup>4</sup>

These considerations also provide the key to understanding the formal fallacy committed by the dialectical syllogism underlying the antinomial conflict.<sup>5</sup> Once again, Kant characterizes the fallacy as a *sophisma figurae dictionis*, or fallacy of equivocation (A499 / B528).<sup>6</sup> And, once again, it reflects a substantive metaphysical error. Kant initially describes this error as an "amphiboly" by which an idea is taken for "a putative representation of something given empirically, and thus of an object to be cognized in accordance with the laws of experience" (A484 / B512). Later, in connection with the discussion of the legitimate



regulative function of the cosmological ideas, he refers to it as a “transcendental subreption,” through which objective reality is ascribed to an idea that merely serves as a rule” (A509 / B537). Consequently, as in the Paralogisms, the fundamental error consists in attributing objective reality to a mere idea, that is, in hypostatizing it. The difference is only that here the product of this hypostatization is thought to be something sensible.

This underlying metaphysical error, which is the immediate consequence of being taken in by the illusory major premise, leads, in turn, to the formal fallacy in the cosmological syllogism. The latter consists in an equivocation regarding the key term *conditioned*, which, Kant claims, is taken in different senses in the two premises. Whereas in the major premise it is taken in the transcendental sense of a pure category, in the minor it is understood in the empirical sense as “a concept of the understanding applied to mere appearances” (A499 / B527). Taking it in the first sense involves the usual abstraction from the spatiotemporal conditions through which the conditioned is given empirically, thereby effectively viewing its relation to its conditions in terms of the model of a conclusion to its premises. Conversely, taking it in the second sense, which is required by its application to “objects of the senses” in the minor premise, necessitates factoring in these very conditions. Thus, the cosmological syllogism (like its psychological counterpart) is formally fallacious. Here, however, rather than merely four fallacious syllogisms involving a single idea (the soul), there are four distinct cosmological ideas, each of which consists in a certain way of thinking the sensible world as a whole.

## B. The System of Cosmological Ideas

As one would expect, Kant derives the specific cosmological ideas from the table of categories. The principle underlying this derivation is that pure and transcendental concepts can arise only from the understanding and, therefore, “that reason really cannot generate any concept at all, but can at most only free a concept of the understanding from the unavoidable limitations of a possible experience, and thus seek to extend it beyond the boundaries of the empirical” (A408–09 / B435–36). Accordingly, Kant claims that “the transcendental ideas will really be nothing except categories extended to the unconditioned” (A409 / B436).

Although this claim may seem surprising in view of Kant’s sharp separation between reason and understanding and his insistence that the former has principles of its own distinct from the latter, there is no inconsistency in Kant’s po-

sition. We have already seen that what reason actually adds to the understanding is the idea of the unconditioned and its associated principle ( $P_2$ ), which is precisely the point Kant is now making.

The question thus becomes which categories are capable of such an extension with respect to the thought of the totality of appearances. Since this extension occurs by means of connecting a category with  $P_2$ , this is equivalent to the question of which categories can provide a rule of synthesis in which the items synthesized are subordinated to one another as conditioned to condition (A409 / B436). As before, only a regressive series or synthesis from conditioned to condition will do, because it alone requires the thought of the totality of appearances demanded by reason. Given these parameters, Kant concludes that the categories of quantity, reality, causality, and necessity qualify as such rules of synthesis and, therefore, are potential sources of specific cosmological ideas.

Since reason, under the direction of  $P_2$ , is naturally led to regard the sensible world as having an extensive quantity or magnitude, the category (or categories) of quantity is clearly involved. Moreover, since this "quantification" concerns the relation of the world to both space and time, which Kant characterizes as "the two original *quanta* of all intuition" (A411 / B438), it involves both the age and the size of the world.<sup>7</sup> But rather than arguing directly for this claim, which he perhaps viewed as self-evident, Kant focuses on the salient differences and similarities in the quantification of time and space.

In virtue of its successive nature, the case of time is clear cut. The key is that we naturally regard the present moment of time as "conditioned" by past moments in the sense that it is conceived as the result of a completed synthesis of these past moments. Kant notes, however, that this applies only to past and not to future times. With regard to the latter, there is no role for a transcendental idea, because there is no thought of completeness or absolute totality (A411–12 / B438–39). In short, the future, unlike the past, is open ended.

Space, by contrast, must be thought as an aggregate (of spaces) rather than a series, since its parts coexist. Consequently, one cannot in the same sense view a particular part of space as the "condition" of another part, which seemingly precludes the kind of regressive synthesis operative in the determination of past time. Nevertheless, Kant suggests that something analogous is involved in the *measurement* of space; for each part of space is limited by other parts, and in this sense it presupposes these other parts as "the condition of the boundaries of the previous ones" (A412 / B439). Thus, the measurement of a determinate space (and of appearances in space) necessarily involves a synthesis or regress from the conditioned to its condition, which is governed by the concept of extensive

magnitude under the direction of  $P_2$ . As Kant puts it, "I can ask about the absolute totality of appearance in space as well in past time" (A413 / B440). He also intimates, however, that the real question is whether *either* inquiry is legitimate.

The categories of quality are involved in the thought of an absolute totality of appearances because "reality in space" or that which occupies space, namely matter, is likewise thought as something conditioned. Here the conditions are the parts of which matter is composed, which, in turn, have their parts or conditions, and so forth. Thus, in virtue of  $P_2$ , we again have a regressive synthesis, the absolute totality or completion of which is demanded by reason (A413 / B440).

Among the categories of relation, Kant maintains that only causality is suitable for the formation of a cosmological idea. As the relation of ground and consequent, its suitability is evident, since the move from consequent to ground or effect to cause is a regress from something conditioned to its condition. Accordingly, it opens up the prospect of an absolute totality of these causal conditions. What may seem problematic, however, is the exclusion of the other relational categories, substance and community, which Kant had already linked to the transcendental ideas of the soul and God. The worry is that, by denying these categories any connection with  $P_2$ , Kant is now flatly contradicting his earlier analysis on which the entire structure of the Dialectic is based.<sup>8</sup>

As is the case in the generation of the transcendental ideas, however, there is no contradiction. The crucial point is that the present concern is with the totality of *appearances*, which is not the case with either the psychological or theological idea. Accordingly, the question is not which categories are capable of yielding transcendental ideas by being extended to the unconditioned; it is rather which categories are involved in the thought of an absolute totality of *appearances*. Otherwise expressed, the project is to determine the different ways in which the spatiotemporal world can be thought as a whole, not the ways in which the unconditioned in general can be thought. Moreover, neither substance nor community qualifies in this regard, since they do not yield a serial ordering of appearances.

Kant is considerably more hard pressed to find a cosmological idea connected with the modal categories. In order to do so he introduces the concept of the contingent in existence, which he claims must be viewed as conditioned and, as such, refers to a condition under which it is necessary. But since the latter is itself contingent with respect to its condition, Kant suggests that we are once again led by  $P_2$  to a regress culminating in the thought of something un-

conditioned, which here means something whose existence is unconditionally necessary (A415 / BB442). Thus, the fourth cosmological idea is that of an unconditionally or absolutely necessary being, that is, a being whose existence is not dependent on that of any other being.

Here, perhaps more than anywhere else in Kant's treatment of the Antinomy, the familiar charge of excessive concern with the architectonic seems to have some force. The basic objection, which once again can be traced back to Schopenhauer, is that the Fourth Antinomy based on this cosmological idea is nothing more than a replay of the Third.<sup>9</sup> Against this it can be pointed out that the Third Antinomy deals with unconditioned *causality* and the Fourth with unconditioned *existence*. Although this is true and suffices to differentiate these cosmological ideas and their corresponding antinomies, it creates the fresh problem of distinguishing the idea of a necessarily existing being pertaining to rational cosmology from the idea of such a being that belongs to rational theology proper. We shall see in the next chapter that Kant does draw such a distinction, but his attempt to carve out space for the concept of a necessarily existing being as a *cosmological* idea, that is, one connected with the thought of the totality of appearances, remains deeply problematic.<sup>10</sup>

Setting that issue aside, however, what is truly important is Kant's claim that each cosmological idea generates two equally compelling but contradictory conceptions of the unconditioned. This would be significant even if it turned out that there were only one such idea, since that would be sufficient to generate the contradiction of reason with itself in its endemic effort to think appearances as constituting a totality. Consequently, it should suffice to analyze two of the four antinomies specified by Kant: the First because it is the most widely discussed and most clearly fits the antinomial picture; and the Third because of its intrinsic importance to Kant's entire philosophy, both theoretical and practical. But before turning to these it will prove useful to consider briefly Kant's division of the cosmological ideas and the corresponding antinomies into two distinct types, each with its own mode of resolution.

### **C. World and Nature and the Distinction between the Mathematical and Dynamical Antinomies**

Near the end of his preliminary discussion of the cosmological ideas, Kant introduces the distinction between the concepts of world and nature. The world is said to signify "the mathematical whole of all appearances and the totality of their synthesis in the great as well as in the small, i.e., in their progress through

composition as well as through division"; nature refers to this very same world, though considered as a "dynamic whole" (A418 / B446). This anticipates and lays the groundwork for the contrast that Kant later draws between Mathematical and Dynamical Antinomies, which mirrors a similar distinction already drawn at the levels of category and principle (A528–32 / B556–60). Here it amounts to a distinction between two kinds of whole: one (world) is a whole in the sense of a complete collection of elements or parts; the other (nature) is an explanatory whole, that is, one in which there are no explanatory gaps. Appealing to the language of the later distinction, Kant characterizes them as, respectively, a "mathematical" and a "dynamic whole" (A418–19 / B446–47).

The key to the contrast between these two kinds of whole lies in the distinction between two kinds of synthesis through which they are thought. In the thought of a *world*, the elements synthesized or collected are necessarily homogeneous in the sense that they are considered merely qua spatiotemporal, that is, as appearances. By contrast, in the thought of *nature* as a whole, the items are linked dynamically, which means that they may be heterogeneous in the sense of not necessarily being all spatiotemporal entities or events (A528–29 / B556–57). Although the latter may seem strange in light of the fact that all the cosmological ideas are supposed to concern the thought of the *sensible* world, it is a consequence of the concept of an explanatory whole, since the latter allows for the possibility that the required closure might necessitate conditions that are not themselves sensible. In short, it leaves conceptual space for "intellectual beginnings" or unconditioned conditions, which have no place in the concept of the world narrowly conceived.<sup>11</sup>

Given this distinction, it follows that the antinomies generated by these two kinds of cosmological idea are subject to significantly different analyses and resolutions. The resolution of the first two or "Mathematical Antinomies," which are concerned, respectively, with the extensive magnitude of the world (its age and size) and the divisibility of matter (what occupies space), turns on showing that both the thesis and antithesis positions are false, since they both rest on an underlying and self-contradictory conception of the sensible world as a self-existent whole. Conversely, since the "Dynamical Antinomies," which deal with the synthesis of heterogeneous elements, do not rest on such a self-contradictory conception, they leave room for the possibility that (properly understood) both sides may be correct: the thesis in affirming the necessity of an unconditioned condition of appearances lying outside appearances; the antithesis in denying that anything unconditioned may be found within the spatiotemporal world.<sup>12</sup>

## II. THE FIRST ANTINOMY

The thesis of the First Antinomy asserts that the world has both a beginning in time and a limit in space. The antithesis denies each of these claims and contends instead that the world is infinite with respect to both time and space. Taking  $x$  to refer to the world and  $F$  and  $I$  to refer to the finitistic and infinitistic positions, respectively, the common assumption underlying this dispute can be symbolized as  $(\exists x) (Fx \vee Ix)$ .<sup>13</sup>

It must be emphasized at the outset that the dispute concerns the nature of the *relationship* between the world and space and time, not, as is sometimes assumed, the nature of space and time themselves.<sup>14</sup> Consequently, the concept of a world, specifically a spatiotemporal world or totality of appearances, is central to the whole analysis and is presupposed by both parties. At issue is whether such a world is limited by or coextensive with its spatiotemporal framework. Although this is an idealized dispute, it has obvious and frequently noted affinities with the actual dispute between Newton and Leibniz.<sup>15</sup>

### A. The Thesis

Like all the proofs used in the Antinomy, that of the thesis of the First Antinomy is apagogic. It purports to demonstrate that the world must be finite in the relevant respects by showing that it cannot be infinite in these same respects. The temporal portion of the argument, on which we shall focus here, turns largely on Kant's analysis of the notion of an infinite series and on the question of its compatibility with the concept of a world. It can be broken down into the following six steps:

1. Assume the opposite: the world has no beginning in time.
2. It follows from this that up to any given moment (the present) an eternity has elapsed.
3. This, in turn, means that an infinite number of successive changes in the states of things (an infinite number of successive events) has actually occurred, that is, an infinite series has been completed.
4. But, according to the "true transcendental concept," the infinity of a series consists in the fact that it can never be completed through a successive synthesis.
5. The concept of an infinite world-series that has "passed away" (been completed) is, therefore, self-contradictory.
6. Consequently, there must be a beginning of the world in time, that is, a first event. (A426 / B454)

The central claim is that the claim that the world has no beginning in time logically requires one to hold that at any given moment of time an eternity has “elapsed” (*abgelaufen*). Since the dispute concerns the sequence of things in time and not time itself, this is taken to mean that at any given moment “an infinite series of successive states of things in the world, each following another, has passed away [*verflossen*].” Presumably, this entails that an infinite series has been completed. But, the argument continues, “the infinity of a series consists precisely in the fact that it cannot be completed through a successive synthesis.” In the remark on the thesis, Kant characterizes this as “the true (transcendental) concept of infinity” (A432 / B460). In light of this concept, the argument concludes: “Therefore an infinitely elapsed world-series is impossible, so a beginning of the world is a necessary condition of its existence.”

*Some Standard Criticisms.* Given that this argument is extremely cryptic and apparently appeals to arbitrary assumptions, it is not surprising that it has not met with much favor. The most prevalent criticism is that the argument involves a rather crude form of psychologism (or, alternatively, that it presupposes transcendental idealism).<sup>16</sup> As Kemp Smith succinctly puts it, “From a subjective impossibility of apprehension he [Kant] infers an objective impossibility of existence.”<sup>17</sup>

Much the same point is also made by Russell, who combines it with an appeal to the Cantorian conception of infinite or transfinite number. In light of this conception, he dismisses Kant’s characterization of the infinitude of a series as involving “the impossibility of the completion of a successive synthesis.” Russell contends, and he has here been followed by many others, that reference to synthesis, which presumably presupposes the mental activity of synthesizing, is out of place in a discussion of the concept of infinity. By including it, Kant succeeds only in introducing, in Russell’s words, “more or less surreptitiously, that reference to mind by which all Kant’s philosophy was infected.”<sup>18</sup> But the notion of infinity (like that of all numbers) refers primarily to a property of classes and is only derivatively applicable to a series. Moreover, Russell points out, “classes which are infinite are given all at once by the defining property of their members, so that there is no question of ‘completion’ or of ‘successive synthesis’.”<sup>19</sup> With this analysis, Kant’s objections against the supposed infinity of the world are swept aside.

Almost as an afterthought, Russell introduces a second, independent criticism. According to this new objection, Kant’s argument would be hopeless, even if one were to allow his talk about successive synthesis, because:

When Kant says that an infinite series can “never” be completed by successive synthesis, all that he has even conceivably a right to say is that it cannot be completed in a finite time. Thus what he really proves is, at most, that if the world had no beginning, it must have already existed for an infinite time. This, however, is a very poor conclusion, by no means suitable for his purposes.<sup>20</sup>

This line of criticism, which has been reiterated by Strawson,<sup>21</sup> amounts to the charge that Kant is guilty of a *petitio principii*. The reasoning goes roughly as follows: either (1) the world has a first beginning in time, or (2) it has existed for all time. If, as the argument assumes, the synthesis cannot be completed (in a finite time), it follows that the world could not have had a first beginning. But since the whole purpose of the argument is to show what follows from the denial of the first conclusion, the proper conclusion is the second. Kant only avoids this consequence by stating that the second is impossible, which begs the whole question.

A somewhat similar criticism was raised by G. E. Moore, whose critique of Kant's argument is part of his attack on idealism in all its forms. Moore's position is that if the argument proves anything at all, it proves that time does not exist (which is hardly the conclusion Kant drew from it). Moore acknowledges that such a result would follow if Kant could really prove that neither the thesis nor the antithesis was true. Moreover, since he accepts the claim of the antithesis that time and space cannot have first moments or parts, he presents the issue in such a way that everything turns on the proof of the thesis. This proof, however, is summarily dismissed on the grounds that it is a “pure fallacy based on an ambiguity in the notion of end.”<sup>22</sup>

Moore's contention, which is reiterated by Bennett, is that Kant confused the true proposition that an infinite series does not have two ends—that is, is not bounded at both ends—with the manifestly false claim that it cannot have one end—that is, that it cannot be bounded at all.<sup>23</sup> Bennett cites the series of natural numbers as an example of a series that is bounded at one end and yet is infinite. In light of this, he claims that Kant's mistake was to infer falsely from the fact that the temporal series has one end (the present moment) that it cannot be infinite. The most, however, that can be inferred from this is that, if infinite, the series could not have had a beginning, which is precisely what the infinitistic position asserts. From this he concludes: “It is, therefore, a pure fallacy to suppose that there cannot have been an infinite series of past hours, simply because that series has an end in one direction and has come to an end now; all that we mean by calling it infinite is that it has no end in the other direction or, in other words, no beginning.”<sup>24</sup>




*Response to These Criticisms.* Keeping Russell's critique in mind, let us consider some of Kant's remarks about the infinite. To begin with, Kant is careful to distinguish his "true (transcendental) concept of infinity," according to which "the successive synthesis of unity [*Einheit*] in the traversal of a quantum can never be completed" (A432 / B460), from what he terms "a defective concept of the infinity of a given magnitude" (A430 / B548). The latter is simply that of a maximum or greatest number. Since there can be no greatest number, this conception enables one to gain an easy but spurious victory over the infinitistic position. Kant's concern is, therefore, to distinguish his argument from one based on that conception.<sup>25</sup> The important point here is that, according to the Kantian definition, the notion of the infinite is not incoherent, which allows the infinitistic position at least to get off the ground.

In a footnote Kant appended to his characterization of the infinite, he remarks that such an infinite quantum "contains a multiplicity [*Menge*] (of given units) that is greater than any number," which he identifies as the "mathematical concept of the infinite" (A432 / B460n.). Assuming that by 'number' (*Zahl*) Kant means natural number, this can be taken as a schematized version of the transcendental concept. It contains a specific reference to number, the schema of quantity (A142 / B182), and it expresses in numerical terms what the "transcendental" or "pure" concept expresses in strictly conceptual terms, namely, the thought of the incompletability or inexhaustibility of the enumerative process.<sup>26</sup> According to this conception, then, to say that a set contains infinitely many members is just to say that however many of its members one has picked out or enumerated, there are still more to count.<sup>27</sup> This is compatible with Russell's characterization of infinite classes as "given all at once by the defining property of their members," and presumably with the Cantorian conception as well.<sup>28</sup>

Also crucial for understanding the argument is the distinction drawn in the remark on the thesis of the Second Antinomy between a *compositum* and a *totum* (A438 / B467). Elsewhere Kant refers to these as a *totum syntheticum* and a *totum analyticum*, respectively.<sup>29</sup> A totum syntheticum is a whole composed of parts that are given separately (at least in thought). Not only does the concept of such a whole presuppose its distinct, pre-given parts, it is also conceived as the product of the collection (in Kant's term, "synthesis") of these parts. Consequently, the question of whether a particular totum syntheticum is possible is equivalent to the question of whether a complete collection of its parts is conceivable. A totum analyticum, by contrast, is a whole, the parts of which are only possible with reference to that whole. Space and time, according to Kant,




are such *tota analytica*, which is why they are characterized as infinite.  At the material universe, the world in space and time, is conceived as a *totum syntheticum*.

Given this characterization of the material universe, it is clear that the alleged contradiction in the infinitistic position must be located in its application of the concept of the infinite, which is itself perfectly legitimate, to the material universe. Since the latter is conceived as a *totum syntheticum*, the thought of the complete enumeration or “synthesis” of its parts, which is built into this concept, contradicts the thought of inexhaustibility, which is similarly built into the concept of the infinite. Kant makes this explicit at the end of his remark on the thesis when he writes:

Now since this synthesis has to constitute a series that is never to be completed, one can never think a totality prior to it and thus also through it. For in this case the concept of the totality itself is the representation of a completed synthesis of the parts, and this completion, hence also its concept, is impossible. [A433 / B461]<sup>30</sup>

The above analysis suggests the nature of the Kantian response to the general charge of subjectivism or psychologism, because it shows that the critique of the infinitistic position turns on a conceptual claim and has nothing to do with the presumed psychological impossibility of grasping or comprehending the infinite.<sup>31</sup> The point can be clarified by noting that the concept of a *totum syntheticum* is here operationally defined in terms of the intellectual procedure through which it is conceived. Consequently, the problem is that the rule or procedure for thinking a *totum syntheticum* clashes with the one for thinking an infinite quantity. The former demands precisely what the latter precludes, namely, completability (at least in principle). In short, we are given two incompatible rules for thinking the same object, one stemming from reason ( $P_2$ ) and the other from the understanding, which amounts to a genuine contradiction.

Similar considerations apply to the Russell-Strawson and Moore-Bennett objections. First, in response to the former’s *petitio* charge, it must be noted that the assumption that the series is infinite entails not merely that it cannot be completed in a finite time but that it cannot be completed at all. Moreover, if  this is the case, then it does not constitute a world (*totum syntheticum*). We thus have two alternatives: either (1) the series does not constitute a world, or (2) there is a first moment. The correct Kantian option is the first; but since the argument presupposes that the series does constitute a world, the proper conclusion is the second.

As we have seen, the Moore-Bennett objection maintains that the argument

commits a “fallacy of ambiguity,” confusing an infinite series, which by definition is open at one end, with a series that has no end at either point. Because of this initial confusion, Kant is said to have reasoned falsely that, since the series does have one end (the present moment), it cannot be infinite. Kant, however, does not claim that a series cannot be infinite if it has one end. As his critique of the “defective concept” of the infinite makes clear, he would have no objection to the infinity of the series of natural numbers. His point is rather that since, as infinite, the series has only one end, it cannot constitute a totality. In other words, the conception of an infinite series that “cannot be completed by successive synthesis” is rejected on the grounds that it violates the totality condition built into the concept of the world as a totum syntheticum.

At this point the critic will likely question the appeal to the concept of a totum syntheticum on which this defense of the thesis of the First Antinomy crucially depends. The objection is that the requirement that an infinite series of past states of the world (or a world extending infinitely in space) be conceived as constituting such a totum seems completely arbitrary. Why, one might ask, can we not simply think of this series as infinite in the sense of being closed at only one end, like the series of natural numbers, without also assuming (*per impossibile*) that it somehow constitutes a “totality”? As Kant himself acknowledges, there is no difficulty in doing this for the series of future states, which can be conceived as infinite.<sup>32</sup> Why, then, should the situation be any different for the series of past states? Granted, in that case the series would not constitute a totum syntheticum, but why should that matter?

This takes us to the heart of the problem. Even though it is not explicitly mentioned, the temporal portion of the thesis argument clearly derives its force from the assumption that past states of the world constitute a totum syntheticum. Moreover, the same can be said, *mutatis mutandis*, of the spatial portion. Thus, the question becomes how this assumption is itself justified; for unless it is, the critique of the infinitistic position mounted in the thesis loses all pretence of credibility.

This justification is to be found in the problematic of the antinomial conflict as a whole rather than in the premises of the thesis argument considered solely in their own terms. Simply put, that the series of past states constitutes a totum syntheticum arises from the application to it of the illusory  $P_2$ . Consequently, what justifies the assumption is precisely its unavoidability. Insofar as one thinks about the sensible world as a whole or, equivalently, about appearances in space and time as constituting a world, one cannot avoid conceiving it as such. In this respect, it is “subjectively necessary.” Correlatively, it is precisely

because one conceives the world in this way that one is constrained to assume that it is either finite or infinite in extent. To raise the very question at issue in the cosmological dispute is already to treat the series as constituting a totum syntheticum. That is why both sides are committed to this assumption. It is also why Kant remarks, "Nothing seems clearer than that between the two, one of whom asserts that the world has a beginning, and the other that it has no beginning but has existed from eternity, one of them has to be right" (A501 / B529). Of course, Kant holds that this "seeming," like the conception of a totum syntheticum of past states on which it is based, is illusory. But before we can evaluate this claim and its connection with transcendental realism, we must consider the argument of the antithesis.

### B. The Antithesis

Like the thesis, the antithesis deals with the world in space and time, not with space and time themselves. Thus, it presupposes the same concept of the sensible world but asserts that this world can have no beginning in time and no limit in space. Also, in accordance with the logic of the antinomial reasoning, it assumes that a demonstration that the world can have no such beginning or limit, that is, that it is not finite, is logically equivalent to a demonstration that it is actually infinite. As before, we shall concern ourselves mainly with the temporal portion of the argument, which can likewise be broken down into six steps:

1. Assume the opposite: the world has a beginning in time.
2. The concept of a (temporal) beginning presupposes a preceding time in which the thing that comes into being does not yet exist.
3. Therefore, the concept of a world coming into being in time presupposes an empty, premundane time.
4. But it is impossible for anything to come into being in empty time because "no part of such a time has . . . any distinguishing condition of its existence rather than its non-existence."
5. Therefore, the "world itself" cannot have a beginning in time.
6. Therefore, the world is infinite with respect to past time. (A427 / B455)

The key steps are 4 and 6. If one assumes that nothing can come into being in empty time, then *a fortiori* the world cannot have had a beginning (in time), since any time prior to the world is by definition "empty." Consequently, the first question is why a beginning in empty time is impossible. The argument turns on the presumed impossibility of empty time (or space) serving as a boundary condition of the world. As Kant notes in remarking on the antithe-

sis, "The proof for the infinity of the world-series and of the sum total of the world rests on the fact that in the contrary case an empty time, and likewise an empty space, would have to constitute the boundary of the world" (A431/B459).

In attempting to show why empty time cannot serve in this capacity, the argument takes an epistemological turn, which is not to be confused with verificationism.<sup>33</sup> Although the crucial fourth step poses the problem in ontological terms, claiming that empty time cannot provide any distinguishing conditions of the existence or nonexistence of the world, the real thrust of the argument seems to be that no part of empty time has any features distinguishing it from any other, which recalls Leibniz's appeal to the principle of the identity of discernibles to refute Clarke's Newtonian thesis that the world came into being at a particular moment in an infinite cosmic time.<sup>34</sup> On a strictly epistemological reading, however, this seems merely to yield the conclusion that the world cannot be *known* to have come into being at any particular time rather than the stronger conclusion that it could not have done so, which is supposedly what is required.

Nevertheless, this ignores the transcendently realistic perspective assumed by the rational cosmologist. Once again, what is essential to this perspective is the presumption of a God's-eye view of things as the norm in terms of which the nature and limits of human knowledge are considered. Consequently, for the transcendental realist the inability to find the distinguishing conditions of the existence of the world in empty time means not merely that the human intellect is incapable of determining the time at which the world came into being but that such a determination cannot be made even by "eyes as discerning as those of God," which is tantamount to claiming that it could not have come into being in time at all.

The same considerations serve to elucidate the sixth step of the argument, which consists in a move from the denial of the finitistic position to the assertion of the opposed infinitistic position. Although Kant himself will reject this move as a non sequitur, since it rests on the false assumption that the world must be either finite or infinite, it is not at all problematic from the standpoint of a transcendental realism under the spell of  $P_2$ . From that standpoint, it is the case that the denial of the finitude of the world entails its infinitude, and vice versa.

*Objections and Replies.* There appear to be two major strategies for defending the finitistic position (or something like it) against the attack contained in step

4. Although both take the argument to be essentially verificationist in nature, they may be seen, more generally, as attempts to defend the finitude thesis by separating it from the boundary-condition requirement.

One, suggested by Strawson, involves trying to legitimize the question, Why did the world begin when it did? by construing it as an internal rather than an external question.<sup>35</sup> So construed, it concerns the order or arrangement of the elements within the world, as in, Why *a* before *b*? By contrast, understood as an external question, it involves a reference to some external factor or condition, which might explain why the world began at  $t_1$  rather than  $t_2$ . The operative assumption is that the concern with the world as a whole rules out the possibility of treating it as an external question but does not preclude regarding it as a legitimate internal question. Thus, the question becomes, Why was a given event the first in the series of events constituting the history of the world? This question is meaningful because one can always assume the possibility either of events prior to the designated event or of the actual series of events being ordered in some other way.

The major problem with this move can be demonstrated by means of Keith Donnellan's distinction between a "referential" and an "attributive" use of a definite description.<sup>36</sup> If the definite description 'the first event' is taken referentially as a "rigid designator," then the question, Why did the world begin with *that* event rather than some other? becomes a perfectly legitimate internal question in the manner suggested above. If, however, 'the first event' is taken attributively to denote the earliest event in the series (whatever it may have been), then the only question to be asked is, Why did the first event occur when it did? But this cannot be regarded as an internal question, since it involves reference to an external temporal framework. It is clear, however, that only the attributive use of the expression 'the first event' is relevant to the problematic of the First Antinomy. Consequently, the reinterpretation of the question as internal does not meet the challenge posed by the antithesis argument.

The second strategy, developed by Bennett, attempts to show that the denial that the world has a beginning does not follow from the stated premises. Bennett admits that, though it is possible to give some sense to the notion of an extramundane space (a claim he uses against the spatial portion of the argument), the same probably cannot be done for a premundane time. Nevertheless, he denies that this justifies the conclusion actually drawn in the argument for the antithesis. As Bennett puts it, "From the impossibility of a premundane time I do not infer the impossibility of a first event. Rather, I infer that if there was a first

event, it occurred at the first time.”<sup>37</sup> Moreover, he continues, “From the true premise that ‘empty time prior to the world’ is a ‘non-entity’, he [Kant] immediately infers not that the first event must have occurred at the first time but rather that there cannot have been a first event.”<sup>38</sup>

Given this analysis, Bennett’s task is to defend the coherence of the assumption that the first event occurred at the first time. This requires defending the intelligibility not only of the notions of a first time and a first event taken individually but also of their conjunction. Although Bennett succeeds in demonstrating the coherence of a first time and the same can be done for a first event, he fails with the third, which is the decisive factor. In other words, while the notions of a first time and a first event are themselves perfectly coherent, the same cannot be said of the notion of a first event at a first time.

In order to conceive of a first time, Bennett suggests that we take as our point of reference any historical event  $H$  and assume that  $n$  represents the number of years from  $H$  back to the first time. The phrase ‘ $n$  years before  $H$ ’ thus designates the first time.<sup>39</sup> The obvious problem suggested by this analysis concerns the possibility of conceiving of times more than  $n$  years before  $H$  and, therefore, before the “first time.” If  $n$  has a finite value (as is required by the argument), then it must be possible to conceive of such times. Bennett’s solution is to contend that any phrase of the form ‘ $K$  years before  $H$ ,’ where  $K > n$ , “makes sense but does not refer to any time.”<sup>40</sup> The point appears to be that we can easily imagine times more distant from  $H$  than  $n$ . Thus, even though ‘ $K$  years before  $H$ ’ does not in fact refer to a time (since *ex hypothesi*  $n$  is the first time), it conceivably could do so, and this is enough to guarantee the coherence of the notion of a first time. This analysis seems perfectly acceptable.

Although Bennett does not specifically defend the notion of a first event, this can easily be done. Given the definition of ‘event’ (*Begebenheit, Ereignis*) as a change of state or alteration of a thing in time, the phrase ‘the first event’ designates the earliest such change to have occurred in the universe. Conceivable events prior to this can be dealt with in the same way that Bennett deals with conceivable times prior to the first time.

As already indicated, however, the difficulty concerns the location of a putative first event at a putative first time. The problem is that, since by ‘event’ is meant a change in the state of a thing, every event presupposes a prior time in which the thing existed in a different state. Consequently, the notion of a first event at a first time, that is, an event not preceded by a time in which the world (the “thing” in question here) was in a different state, turns out to be incoher-

ent. But a beginning of the world at the first time would, on Bennett's analysis, be just such an "event." Its possibility can, therefore, be rejected on the same grounds.

The obvious retort is to grant that a putative beginning of the world is not an 'event' in the above-mentioned sense but to insist that this does not rule out its conceivability. This is correct and allows one to defend the conceivability of a creation or first beginning of the world. The problem, however, is that it does not allow one to contend that this first beginning occurred at the first (or any other) time. The main point, which was discussed in chapter 9, is that quite apart from the question of whether a change is termed an 'event', it is a condition of the possibility of conceiving of the change of a thing in time that we are able to contrast the state of a thing at an earlier time with its state at a later time. The upshot of the matter, then, is that one can perhaps hold with Augustine and many others (including Leibniz) that time began with creation, but one cannot meaningfully claim that creation occurred at the first time.<sup>41</sup> It is, however, the latter that must be established in order to refute the argument under consideration.

As was the case with the thesis, this defense of the antithesis against these objections obviously does not amount to a vindication of the argument itself. In particular, one may be left with the worry that the antithesis shows merely (or at most) that, on the assumption that the world had a beginning, the question why it began at one time rather than another is unanswerable, which seems quite distinct from showing (as claimed) that it could not have begun in time.<sup>42</sup> Once again, however, it is crucial to keep in mind the transcendently realistic standpoint from which the argument is advanced. We have seen that, *from that standpoint*, it does make perfect sense to argue that the impossibility of a beginning of the world in time follows from the impossibility of determining (from a God's-eye view) the time at which it began.<sup>43</sup>

### III. THE THIRD ANTINOMY

We have seen that the Dynamical Antinomies are concerned with "nature," understood as the totality of appearances considered as an explanatory whole. In the case of the Third Antinomy, the relation between conditioned and condition, which supposedly constitutes such a whole, is causal. Moreover, both parties to the dispute assume the validity within experience of "causality in accordance with laws of nature," that is, the mode of causality affirmed in the Second Analogy. At issue is whether it is also necessary, or even permissible, to appeal to



another conception of causality, transcendental freedom, defined as the “faculty [*Vermögen*] of beginning a state from itself [*von selbst*]” (A533 / B561) in order to account for the origin of the world as a whole. Consequently, the issue as defined is explicitly cosmological. Indeed, Kant underscores this in his remark on the thesis, when in confirmation of the “need of reason” to assume such a first beginning, he notes that, with the exception of the Epicureans, all the schools of ancient philosophy affirmed the necessity of positing a prime mover for the explanation of motions in the world (A450 / B478). This exception is significant, given Kant’s identification of the standpoint of the antithesis in all the Antinomies with that of Epicurus.

Kant also claims in this remark that the conception of absolute spontaneity of a cause has important implications for the traditional problem of free will (practical freedom), because it constitutes the major difficulty in conceiving of such freedom. He further suggests that by distinguishing between a “beginning in time” and a “beginning in causality” it becomes possible to allow that certain series of appearances within the world, namely, voluntary human actions and their causal consequences, have an “absolutely first beginning” in the latter sense, even though they are not beginnings in the former sense (A450 / B478). But these issues are beyond our present purview, which is limited to the cosmological dimension of the Third Antinomy.<sup>44</sup>

In affirming the need to posit a first beginning or unconditioned causality in order to account for the origin of the world as a whole, the thesis expresses the demand of reason for explanatory completeness. The antithesis denies the legitimacy of positing such causality, as well as its extension to certain series of appearances within the world, which is not explicitly affirmed in the thesis argument itself. It affirms instead that all causality must be of the Second Analogy type and, in so doing, it expresses the requirement of the understanding to remain within the domain of possible experience. Although it is not immediately clear from Kant’s formulation, the demonstrations are once again intended to be apogogic in nature: each side attempts to establish its case by demonstrating the impossibility of the alternative. In order to capture this, the thesis and antithesis may be taken as arguing, respectively, that not everything takes place according to natural causality and that everything does take place according to such causality.<sup>45</sup>

### A. The Thesis

As already indicated, the thesis advances a claim about the requirements for an adequate explanation of nature, considered as a dynamical whole. It asserts

that, in addition to “causality in accordance with laws of nature,” operative within nature, it is necessary to assume a “causality of freedom” in order to account for such a whole. The argument can be broken down into the following eight steps:

1. The assumption of the opposing view: “there is no other causality than that in accordance with laws of nature.”
2. This means that “everything that happens presupposes a preceding state, upon which it follows without exception according to a rule.”
3. But this entails (by universalization) that the preceding state must itself have come into existence in time. If this be denied, that is, assuming that the preceding state has always existed, then its consequence (the succeeding state) would likewise have always existed. But this contradicts the assumption that the latter has come into existence in time.
4. Since the “causality of the cause through which something happens is always something that has happened,” it presupposes, “according to the law of nature [*nach dem Gesetz der Natur*]” its own antecedent cause, and so forth.
5. Consequently, on the assumption that “everything happens according to mere laws of nature [*nach blossen Gesetzen der Natur*],” there will always be only “a subordinate [*subalternen*] but never a first beginning, and thus no completeness of the series on the side of the causes descending from one another.”
6. “But now the law of nature [*Gesetz der Natur*] consists just in this, that nothing happens without a cause sufficiently determined *a priori*.”
7. Thus, when “taken in its unlimited universality,” the claim that “all causality is possible only in accordance with laws of nature [*nach Naturgesetzen*]” contradicts itself.”
8. Consequently, it cannot be regarded as the sole kind of causality. (A444–45 / B472–74)

Strictly speaking, this argument establishes at most the negative conclusion that causality according to laws of nature is not the only kind of causality, which might seem to be quite distinct from establishing a positive thesis about a non-natural kind of causality involving absolute spontaneity or, equivalently, transcendental freedom. As Kant makes clear in his discussion of the resolution of the Antinomy, however, he is treating these two conceptions of causality as contradictories (A533 / B561). Moreover, insofar as freedom is here understood in a purely negative sense as independence of the conditions of nature, there is

nothing particularly problematic in the move from the insufficiency of the former to the necessity of the latter.

Judging by its treatment in the secondary literature, the same cannot be said about the basic argument against the denial of any non-natural causality. This argument turns on two closely related claims: (1) that natural causality alone can never yield completeness in a causal series (step 5); and (2) the claim that this is the only kind of causality is self-contradictory, when considered as strictly universal (step 7). The former seems to be either irrelevant or question begging: Why should this lack of completeness matter? The latter seems deeply mysterious and is sometimes claimed to rest on a gross conflation.

Neither of these, however, is true. To begin with, the completeness requirement alluded to in step 5 is the underlying presupposition of the whole dispute. Thus, the question is not *whether* nature is to be viewed as a dynamical whole (the completeness requirement) but, assuming that it must be, how such a whole is to be conceived. Consequently, by appealing to the requirement at this point, the thesis argument is neither introducing extraneous considerations nor begging the question at hand.

The charge of a self-contradiction is more complex and requires a detailed discussion. It is deduced in step 7 from the explanatory incompleteness affirmed in step 5, by means of the principle that “nothing happens without a cause sufficiently determined *a priori*,” which is presented in step 6 as “the law of nature.” Accordingly, step 6 is the nerve of the argument.

Given its obvious centrality, it is not surprising that the nature of this law and its relation to the “laws of nature” referred to elsewhere in the argument have been the focal point of both the criticism of this argument and the attempts to interpret it. Problems begin with the interpretation of “a cause sufficiently determined *a priori*.” Perhaps the most natural reading of this phrase is to take it as referring to a sufficient cause as contrasted with a partial cause, that is, a set of necessary and jointly sufficient conditions. Although the Second Analogy purports to show that every occurrence must have a sufficient cause in this sense, this, of itself, fails to yield the alleged contradiction. As Schopenhauer pointed out, it is one thing to assert that every event requires a cause sufficient to produce the effect and quite another to assert that there must be a completion in the series of antecedent causes leading up to the event.<sup>46</sup> Consequently, on this interpretation, the argument only gets off the ground by conflating these two claims. Moreover, Schopenhauer is once again followed closely by Kemp Smith, who remarks that the question of whether *A* itself is suffi-

ciently explained is distinct from the question of whether *A* (being given) is sufficient to explain *B*.<sup>47</sup>

Fortunately, there is little textual support for the Schopenhauer–Kemp Smith reading. As Bennett notes, the claim is that the cause must be *sufficiently determined*, not that it must be sufficient; and he adds (following Heimsoeth) that “‘*a priori*’ here has its traditional meaning of in advance of or prior to rather than the specifically Kantian sense of ‘independently of experience’.”<sup>48</sup> But rather than attempting to explain what the requirement that a cause be sufficiently determined prior to its occurrence amounts to or how the argument based on this conception is supposed to work, Bennett confesses defeat and offers instead a brief canvass of the supposedly inadequate accounts in the literature.<sup>49</sup>

The most promising of these is a suggestion by Ewing to the effect that if there were only causality of nature, no explanation would be ultimate in the sense of leaving nothing further to be explained.<sup>50</sup> Bennett rejects this proposal, however, on the grounds that it takes the target of the thesis to be a version of the principle of causality that holds *both* that there is only the causality of nature *and* that every event has an ultimate explanation. This is deemed unacceptable because it makes the opponent “into such an obvious straw man that Kant cannot have taken it seriously or supposed that the thesis-arguer would do so.”<sup>51</sup>

What is particularly curious about Bennett’s dismissal of this target as a straw man is that it represents the view of one of his favorite philosophers, namely, Leibniz. In fact, we need only substitute the principle of sufficient reason for what Bennett terms a “principle of causality” in order to arrive at the Leibnizian position as it is articulated in the polemic with Clarke. There and elsewhere, Leibniz maintains that every occurrence has a sufficient reason *both* in the sense that it has an antecedent cause *and* in the sense that it has an ultimate explanation (accessible only to God) based on its role within the total context of the possible world actualized by the divine will.<sup>52</sup>

This is evidenced by the fact that there Leibniz construes the principle of sufficient reason as applying equally “to any thing’s existing . . . to any event’s happening . . . to any truth’s taking place.”<sup>53</sup> In Kantian terms, this means that he regards it both as a logical principle requiring adequate grounds for any conclusion and as a real or causal principle governing occurrences or states of affairs. Moreover, if, as Leibniz seems to have done, one regards these requirements as two expressions of a single principle having both logical and metaphysical force, then it is natural to take it as requiring an ultimate explanation

or grounding for any *explanandum* analogous to the logical grounding in principles required for the justification of conclusions. In short, in virtue of his understanding of the principle of sufficient reason, Leibniz is committed to the view that there must be both an antecedent cause and a complete explanation (at least in principle, or for God) for every occurrence or state of affairs, including the existence of the world as a whole.

Thus, rather than attacking a straw man, as Bennett suggests, the thesis of the Third Antinomy can be taken as providing a *reductio* of Leibniz's position, particularly as formulated in his polemic with Clarke regarding the nature and scope of the principle of sufficient reason.<sup>54</sup> The basic point of this *reductio* is that the two requirements built into what purports to be a single principle pull in opposite directions. Specifically, the requirement of explanatory completeness affirmed in step 5 cannot be reconciled with the principle that *all* causality and, therefore, all explanation be in accordance with laws of nature, because the latter mode of explanation can never produce the requisite explanatory completeness.<sup>55</sup> Consequently, when this principle is "taken in its unlimited universality," as it must be when applied to nature as a whole, it yields the contradiction referred to in step 7. Leibniz cannot have it both ways.

Nevertheless, it would be a mistake to see the thesis argument as nothing more than a *reductio* of the Leibnizian position, or even of this particular conception of the principle of sufficient reason, since that would mean that the conclusion could be avoided simply by jettisoning this dogmatic principle.<sup>56</sup> What blocks this easy dismissal of the argument is that the pivotal step 6 is, at bottom, a thinly disguised expression of  $P_2$ . In other words, the "law of nature" that "nothing happens without a cause sufficiently determined *a priori*" is logically equivalent to the principle that "if the conditioned is given, then the whole series of all conditions for it is also given," when the latter is applied to nature considered as a dynamical whole. Thus, not just Leibniz's but *any* naturalistic conception of causality runs afoul of this principle when "taken in its unlimited universality," because it cannot account for the "givenness" of the series as a whole (the completeness condition).

## B. The Antithesis

As already indicated, the antithesis denies the possibility of a causality of freedom and affirms instead that "everything in the world happens solely in accordance with laws of nature" (A445 / B473). It can be broken down into the following eight steps, which are supplemented by some remarks on the "illusion of freedom."

1. The assumption of the opposing view: there is freedom in the transcendental sense.
2. This means that there is (a) "a faculty of absolutely beginning a state, and (b) a series of its consequences."
3. This entails that (a) a series of occurrences will have its absolute beginning in a spontaneous cause and that (b) this causality itself will have an absolute beginning.
4. Consequently, "nothing precedes it through which this occurring action is determined in accordance with constant laws."
5. But: "Every beginning of action . . . presupposes a state of the not yet acting cause."
6. Moreover, if, as is being assumed, it is not only a "*dynamical* beginning" (the beginning of a causal sequence) but also a *first* beginning, it "presupposes a state that has no causal connection at all with the preceding state of the cause,<sup>57</sup> i.e., in nowise follows from it."
7. "Thus transcendental freedom is contrary to the causal law [*der Kausalgesetze*], and is a connection [*Verbindung*] of the successive states of effective causes in accordance with which no unity of experience is possible, which thus cannot be encountered in any experience."
8. The idea of such freedom is, therefore, "an empty thought-entity" that is, there can be no transcendental freedom. (A445–47 / B473–75)

This argument, which represents the standpoint of a "pure empiricism," is relatively straightforward and lacks any surprising premises, such as step 6 in the thesis argument. After specifying the thesis to be refuted, as well as its salient implications in the first three steps, the argument in step 5 appeals to what appears to be (though it is not presented as such) a basic tenet of the Analogies, to the effect that every change must be connected to the antecedent state of the agent. Step 6 then points out that the conception of a dynamical first beginning, to which the doctrine of transcendental freedom is committed, violates this condition; from which it is inferred in step 7 that the latter cannot be found in experience, since it conflicts with the "causal law" specifying the conditions of the unity of experience. Finally, step 8 infers the absolute from the empirical impossibility of transcendental freedom, dismissing it as "an empty thought-entity."

This argument has generally met with two types of response. On the one hand, there is the view of Schopenhauer that it reflects Kant's critical position and is perfectly satisfactory.<sup>58</sup> A modified version of this view is presented by

Strawson, who describes it as a “simple denial of freedom,” which is consistent with the results of the Second Analogy.<sup>59</sup> On the other hand, there are those critics, who, while likewise regarding the argument as essentially of a piece with the Analytic, reject it for that very reason as question begging or circular. The objection is that it is illegitimate for Kant to appeal to the results of the Analytic at this point, particularly insofar as these results are deemed inseparable from transcendental idealism and this idealism is what is ultimately at stake. In addition, it is sometimes maintained that a radically un-Kantian equation of freedom with lawlessness is needed to make the argument work; and that by appealing to the conditions of the unity of experience, the argument rules out only a spontaneous cause within nature, which leaves untouched the central contention of the thesis that we must assume such a cause outside nature.<sup>60</sup>

With regard to the first objection, it is clear that the antithesis argument neither begs any question by illicitly assuming the validity of the causal principle nor presupposes the truth of transcendental idealism. It is not guilty of the former because the validity of this principle (within experience) is assumed by both parties to the dispute. At issue is merely whether it is also possible to find conceptual space for a different conception of causality (transcendental freedom). Moreover, here the antithesis clearly argues for, rather than simply assumes, its denial. Similarly, it is not guilty of the latter because the dispute between transcendental realism and transcendental idealism does not arise at the empirical level; it arises only when empirically valid principles are extended beyond the limits of experience.

The objection that the argument only works by equating freedom with lawlessness, which supposedly conflicts with Kant’s own views, is likewise beside the point. First, it does not really concern the cosmological side of the dispute, which is the central topic of the Third Antinomy proper. Second, “lawlessness” is here understood in contrast to the “law of nature” or, more generally, causal laws. Accordingly, it has nothing to do with the connection between freedom and moral principles, which is the sphere in which Kant famously denies a lawless freedom.<sup>61</sup> Third, here, as elsewhere, the antithesis represents the standpoint of a pure empiricism rather than Kant’s own critical position.

The objection that the argument moves from the rejection of transcendental freedom within nature to its complete rejection or, equivalently, from its empirical to its absolute impossibility, does, however, strike a nerve and calls for further comment. In dealing with it, must once again keep in mind that the argument represents the transcendentially realistic standpoint of a pure or dogmatic empiricism and that if appearances are regarded as things in themselves,

then what for Kant are merely epistemic conditions are necessarily interpreted as having ontological force. Accordingly, the antithesis argues consistently from its own presuppositions, as Kant himself later notes when he remarks that “if appearances are things in themselves, then freedom cannot be saved” (A536 / B464) and, even more emphatically, that “[i]f we would give in to the deception of transcendental realism, then neither nature nor freedom would be left” (A543 / B571).

#### IV. TRANSCENDENTAL IDEALISM AND THE RESOLUTION OF THE ANTINOMY

This last point takes us to our central concern: the connection between Kant’s analysis of the antinomial conflict and the opposition between transcendental realism and transcendental idealism. The issue can be broken down into two parts. The first, which applies (albeit differentially) to all four of the Antinomies, concerns Kant’s attempt to show that the whole dispute rests on a misunderstanding, stemming from the implicit commitment of all participants to transcendental realism, and can be resolved by opting for transcendental idealism. The second, which applies only to the first two or Mathematical Antinomies, concerns Kant’s indirect proof of transcendental idealism from the allegedly contradictory nature of the entailments of transcendental realism. Whereas the former may be said to treat transcendental idealism as a sufficient condition for the resolution of the Antinomy, the latter effectively elevates it into a necessary condition.

##### A. “Transcendental Idealism as the Key to Solving the Cosmological Dialectic”

The above is the heading for the sixth section of the Antinomy chapter (A490 / B518) and is directly followed by the contrast between transcendental idealism and transcendental realism, which we considered at length in chapter 2. It is only in the subsequent section, however, that Kant applies this contrast to the antinomial conflict. As we have already seen, Kant begins this crucial section by remarking that the entire Antinomy rests on the illusory principle that “[i]f the conditioned is given, then the whole series of all conditions for it is also given,” that is, on  $P_2$ . We have also seen why Kant holds that the hypothetical syllogism, in which this forms the major premise and which underlies the entire antinomial conflict, commits a fallacy of equivocation, which is a direct consequence of being deceived by this illusion. But Kant is well aware that this is not



enough to end the matter, since, regardless of any hidden fallacies, it still seems that one of the parties must be right (A501 / B529). In short, the illusion has not yet lost its grip.

Since, as natural and unavoidable, the illusion cannot be removed completely, the therapeutic function of transcendental idealism in this context is to loosen this grip, that is, to lessen its power to lead to error. As was the case in the Paralogisms and will be again in the Ideal, an essential part of the strategy consists in showing that the deceptive power of the illusion is a function of the metaphysician's implicit commitment to transcendental realism. But since in this case the illusion is two-sided, Kant will also suggest that transcendental idealism provides a fresh way of looking at the situation according to which the conflict between the competing claims, which from the standpoint of transcendental realism seems irreconcilable, may be seen to be merely apparent because it arises from this illusion.<sup>62</sup>

Kant's first step in this therapeutic process is to distinguish between a condition being given (*gegeben*) and the finding of it "set as a task" or "**given to us as a problem**" (*aufgegeben*) (A498 / B526).<sup>63</sup> The task of seeking the conditions for every conditioned stems from the logical maxim of reason or the intellectual categorical imperative ( $P_1$ ).<sup>64</sup> Accordingly, Kant treats the claim that we have such a task as "analytic and beyond any fear of a transcendental criticism" (A498 / B526). It is not, however, analytic that the sought-for conditions are themselves "given" in the sense of being there to be found (even in principle or from a God's-eye view of things). In fact, this is precisely the illusion from whose grip liberation is needed.

Once again, however, this turns out to be beyond the resources of transcendental realism, which serves rather to reinforce the illusion. This is because the illusion in its cosmological guise just is that the totality of conditions (here appearances) exists in a timeless manner as a thing in itself, that is, as a higher-order object for a putative "pure understanding." Consequently, if (as transcendental realism does) one takes appearances, that is, spatiotemporal entities, as things in themselves, then, in following  $P_1$ , it inevitably takes the total series of conditions to be actually given, rather than simply regarding the search for these conditions as set as a task. As Kant here puts it,

If the conditioned as well as its condition are things in themselves, then when the first is given not only is the regress to the second **given as a problem**, but the latter is thereby really already **given** [*gegeben*] along with it; and, because this holds for all members of the series, then the complete series of conditions, and hence the unconditioned is thereby simultaneously given, or rather it is presupposed by the fact that

the conditioned, which is possible only through that series, is given. Here the synthesis of the conditioned with its conditions is a synthesis of the mere understanding, which represents things **as they are** without paying any attention to whether and how we might achieve acquaintance with them. [A498 / B526–27].

The last sentence of this important passage is particularly crucial for understanding Kant's resolution of the Antinomy. To begin with, "a synthesis of mere understanding" is just the "intellectual synthesis" of the B-Deduction. As such, it involves a pure or transcendental use of the categories, which, by abstracting from the sensible conditions of their application (the transcendental schemata), yields merely rules for the thought of "objects in general." Thus, the claim that such a synthesis "represents things as they are" should not be taken to suggest that it provides knowledge of things as they are in themselves. It is rather that transcendental realism is constrained to take it in this way, since it ignores the question of "whether or how we might achieve acquaintance with them."

This consequence does not obtain, however, if, with transcendental idealism, we take spatiotemporal objects merely as appearances. For under that assumption, the objects are only given in and through the regress from conditioned to its condition, that is, through an empirical synthesis, which, as such, factors in the spatiotemporal conditions through which the conditions are given. Consequently, transcendental idealism not only allows for but requires precisely what transcendental realism seems forced to deny, namely, a principled distinction between a condition being given and its search merely set as a task (A499 / B527).

Moreover, it follows from this that transcendental idealism is not committed to the illusory assumption that the totality of conditions for any conditioned appearance is already given in the above-mentioned sense. To be sure, it remains "natural" to think of a series of conditions as given in this way and, therefore, as having to be either finite or infinite in the relevant respects. But the transcendental idealist is equipped with a critical tool, making it possible to resist being seduced by this way of thinking.

In the context of the cosmological debate, this tool operates by carving out conceptual space for an alternative that is not recognized by the transcendental realist, namely, that both parties to the dispute are wrong (or, in the case of the Dynamical Antinomies, may be correct). Kant approaches this resolution of the conflict by alluding favorably to Zeno the Eleatic, whose dialectical thought had previously been appealed to by Bayle in support of his attack on the pretensions of reason.<sup>65</sup> Unlike Bayle, however, Kant does not take the

thrust of Zeno's dialectic to be skeptical. Thus, rather than accusing Zeno of wishing to deny two mutually contradictory propositions, for example, the world is finite and the world is infinite, Kant takes him as anticipating his own critical position by questioning the assumption that such propositions are really contradictories. According to Kant's diagnosis, this is done by showing that the two mutually conflicting judgments presuppose an inadmissible condition, which entails that they are not really contradictories and that both are false. As an empirical example, Kant cites the proposition that every body either smells good or smells bad, with the inadmissible condition being that every body has an odor. Since there obviously are bodies totally devoid of odor, both alternatives are false (A502–03 / B530–31).

In the cosmological case, the inadmissible condition is that the world (the totality of appearances) is given as a thing in itself (A504 / B532). As we have seen, because of the underlying illusion, the transcendental realist is in no position to challenge the admissibility of this condition, whereas the transcendental idealist happily is. Thus, Kant can claim that the magnitude of the world is neither finite nor infinite, for the simple reason that it is not the sort of "thing" that is capable of being assigned a magnitude or, more precisely, it is not a thing at all. From the logical point of view, this means that the theses and antitheses of the Mathematical Antinomies are contraries rather than contradictories, which again allows for the possibility (here realized) that both are false. They would be genuine contradictories if the antithesis were merely to claim that the world is not finite; but, as we have seen, it goes beyond this in claiming that it is actually infinite in magnitude.<sup>66</sup>

This diagnosis underscores the systematic significance of the last step in the arguments of the antitheses and reveals the deep error in those interpretations (for example, Schopenhauer's and Strawson's) that equate the position of the antitheses with Kant's own views. As we have seen in the case of the First and Third Antinomies, it is in this step that pure empiricism reveals its transcendently realistic stripes by moving beyond the simple negation of the thesis to the assertion of its own positive alternative view. This is also why Kant suggests that pure empiricism becomes dogmatic and that it, like its dogmatic opponent, "says more than it knows" (A472 / B500).<sup>67</sup>

As we have also seen, however, this diagnosis leads to significantly different resolutions of these two antinomies. This is largely because pure empiricism says more than it knows in different ways in the different types of antinomy. In the case of the First Antinomy, it does so by improperly inferring, from the correct premise that the world cannot have a finite magnitude, that it must have an

infinite one. In the case of the Third Antinomy, it does so by improperly inferring, from the likewise correct premise that all *empirically cognizable causality* must conform to laws of nature, that *all causality* must conform to such laws. From the standpoint of transcendental realism this inference is valid (as is unfortunately the opposed one of the thesis). But it rests on the dogmatic assumption that epistemic conditions are also ontological conditions. Consequently, here transcendental idealism creates the prospect of limiting (rather than simply rejecting) the claim of the antithesis, thus creating conceptual space for the thought of transcendental freedom as a non-empirical mode of causality. Nevertheless, this is all that transcendental idealism can directly contribute to the problem of freedom, since it cannot guarantee that this conceptual space is filled. That is why Kant emphatically denies having established the reality or even the real possibility of freedom, claiming only to have shown “that nature at least does not conflict with causality through freedom” (A558 / B586). As we are about to see, that is also why only the Mathematical Antinomies can enter into Kant’s indirect proof of transcendental idealism.

## B. The Indirect Proof

Kant’s indirect proof is the culmination of his critical resolution of the Antinomy. It purports to show not only that transcendental idealism provides a key to this resolution but also that it is the only key: one either accepts transcendental idealism or falls victim to the “euthanasia of pure reason.” This seems like an exceedingly bold claim, which Kant supports with what appears to be an exceedingly slim line of argumentation. The entire proof is based on a dilemma and proceeds as follows:

If the world is a whole existing in itself, then it is either finite or infinite. Now the first as well as the second alternative is false (according to the proof offered above for the antithesis on the one side and the thesis on the other). Thus it is also false that the world (the sum of all appearances) is a whole existing in itself. From which it follows that appearances in general are nothing outside our representations, which is just what we mean by their transcendental ideality. [A506–07 / B534–35]

The logical form of this argument is a *modus tollens* combined with an immediate inference. The denial of the consequent (“the world is either finite or infinite”) is used, in turn, to deny the antecedent (“the world is a whole existing in itself”), which is then taken to entail the conclusion: “appearances in general are nothing outside our representations.” This latter statement is the thesis of transcendental idealism.

Assuming that the proofs of the thesis and the antithesis are valid, the *modus tollens* is in order.<sup>68</sup> Consequently, the weight of the argument falls on the immediate inference from the negative results of the Mathematical Antinomies. This appears, however, to be quite problematic. In fact, if one confines oneself to the terms of the argument, it seems perfectly possible to accept the conclusion of the *modus tollens*, that the world (the sum of all appearances) is not a whole existing in itself, and to reject the idealistic result that is presumed to follow immediately from it.

Part of the problem can be attributed to the casual way in which this result is expressed. The claim that “appearances in general are nothing outside our representations” might be taken to be analytic, since it can be derived directly from the definition of ‘appearance’, without any reference to the preceding argument. This is only a minor defect, however, which can be avoided simply by reformulating the claim to read: “Spatiotemporal objects are appearances”; as such, they are also “mere representations, which, in the manner in which they are represented, as extended beings, or as series of alterations, have no independent existence outside our thoughts.” But this reformulation also brings into sharper focus the main problem: How can one infer transcendental idealism from the simple denial of the antecedent? The move from the statement “the world (the sum of all appearances) is not a whole existing in itself” to transcendental idealism seems, at least at first glance, to be another example of the kind of gross non sequitur that critics have frequently claimed to find in Kant.

The difficulty can also be put in another way. The proof for transcendental idealism rests ultimately upon the assumption that it provides the only possible basis for avoiding the antinomial conflict. But this assumption might be challenged. For example, framing the issue in the linguistic mode, one could argue that the problem with both the thesis and antithesis positions stems from their construing ‘world’ or ‘the sum of all appearances’ as a naming expression or definite description. Given this assumption, it makes good sense to inquire about the magnitude of the referent of ‘world’ and to presuppose that this magnitude must be either finite or infinite in whatever respects are being considered. This suggests, however, that all that is required to resolve the dispute is to show that ‘world’ does not refer. Indeed, this is surely the most natural reading of the conclusion of the *modus tollens*. Since ‘world’ does not refer to any entity at all, its referent is neither finite nor infinite. To be sure, it follows from the fact that ‘world’ does not refer at all that it does not refer to a thing in itself; but it also follows that it does not refer to the appearances of some unknown thing. In any event, it seems both unwarranted and unnecessary to make the further as-

sumption that the spatiotemporal objects, which when taken collectively were mistakenly supposed to provide the referent of 'world', are themselves mere appearances.<sup>69</sup>

Although this line of criticism seems plausible, it suffers from a failure to consider the essential role of transcendental illusion in Kant's analysis. In the present context, this illusion explains why it seems perfectly natural, even unavoidable, to take 'world', or its equivalents, as a naming expression. Continuing with the linguistic analysis, the term seems to function more like 'army' than, say, 'the average family' in designating an actual collection, indeed, one that is "given" (in an empirical sense) and with respect to which the determination of magnitude seems perfectly in order.<sup>70</sup> To be sure, they differ in that (logically) there can be only one world, whereas there can be any number of armies. Nevertheless, this difference does not seem decisive with regard to the question at issue, since one can draw the same comparison with a particular army.

Expressed in more Kantian terms, the trouble with the concept of the world is that it is pseudo-empirical. As we have already seen, this means that, on the one hand, it appears to signify a high-order empirical object (such as an army), while, on the other hand, since it involves the thought of an absolute completeness of conditions, it exempts itself from the conditions under which alone its purported object could be determined empirically. It is the latter that makes it into an idea of reason and, therefore, merely pseudo-empirical. Moreover, for this very reason, it involves an internal contradiction, which, in turn, explains why it yields contradictory entailments.<sup>71</sup>

Nevertheless, even accepting the correctness of this analysis, which suggests that we cannot simply deny that 'world' is a referring expression without also exposing the underlying illusion that naturally leads us to view it as such, problems remain with the immediate inference to transcendental idealism. This is because the inference is not, as the text suggests, really immediate but is based instead on two suppressed premises. The first is that the antecedent ("the world . . . is a whole existing in itself") is entailed by transcendental realism. The second is that transcendental realism and transcendental idealism are contradictory philosophical standpoints. It follows from the first of these premises that the conclusion of the *modus tollens*, which is the negation of the antecedent, entails the negation of transcendental realism. It follows from the second that this negation is equivalent to the affirmation of transcendental idealism. Given these premises, Kant's overall argument is clearly valid. The main question, therefore, concerns the truth of these additional premises. Since the first was argued for at length in chapter 2, we shall focus here mainly on the second.

Fortunately, in light of what we have already seen, it should be fairly clear that transcendental realism is necessarily committed to the proposition that “the world . . . is a whole existing in itself.” The key points are that  $P_2$ , when applied to appearances, entails that the sum total of these appearances is a whole existing in itself and that transcendental realism, as here understood, is committed to the objective validity of  $P_2$ .

The former is expressed in the hypothetical syllogism underlying the Antinomy as a whole. Recall that the major premise of this syllogism (“If the conditioned is given, then the whole series of all conditions for it is also given”) just is  $P_2$ . Moreover, the conclusion, which Kant does not bother to spell out, is that the whole series of conditions for objects of the senses must be given. But since this whole series constitutes “the world” and since “givenness” in the sense in which it is operative here means givenness for a pure understanding or an intellectual synthesis, which abstracts from the sensible conditions of empirical givenness, the world is taken to be existing “in itself” or, equivalently, as a thing in itself. Consequently, that the world (as defined here) exists in this way is entailed by  $P_2$ .

As we have already seen, similar considerations also account for transcendental realism’s commitment to this illusory principle. Since the consideration of cognition from a putative God’s-eye view is definitional of transcendental realism, it is incapable of recognizing the illusory nature of  $P_2$ . Consequently, it treats this principle as objectively valid. But this, in turn, requires that it also accept its logical consequences, which means that it is committed to the proposition about the world that generates the antinomial conflict.

Given this, the remaining steps of the indirect proof are easy to supply. Since transcendental realism necessarily assumes that the world is a whole existing in itself, it likewise must assume that it is either finite or infinite in the relevant respects. But the analysis of the Mathematical Antinomies (for this purpose the First Antinomy suffices) has shown that this world can be neither finite nor infinite. Consequently, both the conception of the world, which has been shown to be self-contradictory, and the transcendental realism underlying it must be rejected. Finally, given the dichotomy between transcendental realism and transcendental idealism, the negation of the former is logically equivalent to the affirmation of the latter. In short, transcendental idealism is true.

It would be an understatement to say that this argument has not met with a favorable reception. Somewhat surprisingly, in light of his wholesale rejection of Kant’s theory of reason, one commentator who is relatively sympathetic is Walsh. On his reading, the Antinomies are to be taken seriously because they

correspond to positions to which actual metaphysicians subscribe and have their source “in a concept or principle whose internal incoherence is not at first apparent.”<sup>72</sup> Walsh also thinks that this troublesome principle is connected with these metaphysicians’ underlying transcendental realism, but he differs from Kant both in his identification of the nature of this principle and his analysis of its implications.

Following Al-Azm, Walsh identifies the contradiction-generating premise with the principle of sufficient reason, which was accepted (albeit in different senses) by both Leibniz and Clarke. The problem with this principle, according to Walsh, is that it both demands and precludes there being “in the universe some fact, event or existent which is ultimate and self-explanatory.”<sup>73</sup> In other words, there is a built-in ambiguity in this principle, one side of which is exploited by each of the participants in the cosmological dispute.

This clearly constitutes a revision in Kant’s account, since he identified the conflict-generating proposition with the transcendental realist’s assumption that the sensible world exists absolutely or as thing in itself rather than with the principle of sufficient reason. Walsh’s justification for this suggested revision is that it avoids what he takes to be a major flaw in the analysis actually presented in the *Critique*, namely, it commits Kant to the assumption that transcendental realism is not simply false but *necessarily false*, which implies that its contradictory (transcendental idealism) is necessarily true. This, Walsh thinks, is implausible, since “the idea of a world of the senses [as opposed to a world of appearances] existing absolutely seems at first sight quite intelligible.”<sup>74</sup> As Walsh points out, however, this revision comes at some cost, since the location of the root of the antinomial conflict in the ambiguous principle of sufficient reason rather than in an underlying transcendental realism invalidates the proof of idealism. Accordingly, he concludes that “[t]he argument of the Antinomy as here interpreted . . . lends support to the case for transcendental idealism, without amounting to an indirect proof of that doctrine.”<sup>75</sup>

There are two points to be made in response to Walsh’s reconstruction. First, he mistakes the symptom for the disease, which is a consequence of his divorce of the discussion from the underlying theory of reason. Accordingly, Walsh is correct in his diagnosis of the ambiguity built into the principle of sufficient reason and in claiming that each side exploits one side of this ambiguity (this is particularly clear in the Third Antinomy), but he fails to see that each side’s dogmatic appeal to this principle has its roots in their shared attachment to  $P_2$ . In other words, it is precisely because they are in the grip of the underlying illusion that the totality of conditions are given that both the thesis and antithesis



positions appeal uncritically to this principle (albeit under different interpretations) in their endeavors to think the world as a dynamical whole.

Walsh's misunderstanding is revealed in his suggestion that the move "from a realist to an idealist view of the status of things in space and time, allows the transformation of the dogmatic principle of Sufficient Reason into the 'critical' principle of the second Analogy."<sup>76</sup> Although this is clearly intended as a sympathetic effort to capture the "moment of truth" in Kant's analysis of the antinomial conflict, while resisting the conclusion that it amounts to an actual demonstration of transcendental idealism, it involves a significant distortion of Kant's position. The move from transcendental realism to transcendental idealism does not convert a dogmatic principle of *reason* into a 'critical' principle of the *understanding* (that makes no sense for Kant); rather, it makes possible the attribution of merely regulative status to a principle, which from the standpoint of transcendental realism is necessarily viewed as constitutive.<sup>77</sup>

Similar considerations also apply to the second point, which concerns Walsh's claim that, if sound, Kant's argument proves too much, namely, that transcendental realism is necessarily false and transcendental idealism necessarily true. Walsh is certainly correct in stating that the idea of the sensible world as existing absolutely, that is, apart from its relation to the conditions of human sensibility, seems perfectly intelligible. Indeed, it is "intelligible" in the twofold sense of being conceivable and being non-sensible. He is likewise correct in suggesting that, *in a certain sense*, Kant's argument leads to the conclusion that transcendental realism is necessarily false and transcendental idealism necessarily true. Nevertheless, there need not be any incompatibility between these two claims.

To begin with, what makes the proposition that the sensible world exists absolutely seem perfectly intelligible is precisely that it expresses a principle of reason ( $P_2$ ). But it must be kept in mind that this principle is illusory. The problem is that pre-critical philosophers (as transcendental realists) were not equipped to detect this illusion, which means that they proceeded quite reasonably on the basis of their underlying assumption. Moreover, Walsh conveniently glosses over what, for Kant, is clearly the most problematic feature of this conception, namely, its pseudo-empirical nature. As we have already seen, this stems from the fact that it contains the thought of this world as a totality, that is, as something unconditioned, which, as we have also seen, is due to the influence of  $P_2$ .

If this is correct, then what makes transcendental realism ultimately incoherent and, therefore, "necessarily false," only emerges through its combina-

tion with  $P_2$ , since this leads such a realist to regard a putatively *an sich* existing sensible world as a totality (*totum syntheticum*), which, in turn, generates the antinomial conflict. Thus, were it not for this fateful combination, transcendental realism could have continued indefinitely on its parallel dogmatic paths exemplified by Plato and Epicurus, which is precisely what Kant meant by the claim that the discovery of the Antinomy aroused him from his dogmatic slumber. Correlatively, what makes transcendental idealism “necessarily true” is not that the opposed thesis that spatiotemporal objects exist as things in themselves or, more precisely, that spatiotemporal properties and relations pertain to things as they are in themselves, is self-contradictory; it is rather that it provides the only means for dissolving this combination, which, if left intact, issues in the “euthanasia of pure reason.” In short, the truth of transcendental idealism, at least insofar as it concerns the Antinomy, amounts to its therapeutic indispensability, which is precisely what the indirect proof shows.<sup>78</sup>

Against this it will doubtless be argued by some that the above amounts to an “anodyne” interpretation of transcendental idealism, the reduction of it to a supposedly innocent doctrine of “epistemological modesty” or “humility,” which, as such, fails to capture its intended metaphysical force. In particular, Guyer insists that what differentiates Kant’s treatment of the Antinomy in the *Critique* from his earlier treatment of the issues involved is that “methodological restrictions on the claims of sense and reason” are replaced by “metaphysical dogma.”<sup>79</sup> Accordingly, he dismisses Kant’s indirect proof as another of his futile attempts to establish a metaphysical idealism, which, it turns out, is irrelevant to his “transcendental theory of experience.”<sup>80</sup>

Apart from his useful emphasis on the methodological thrust of Kant’s earlier discussions of the issues related to the Antinomy, Guyer is wrong on all counts.<sup>81</sup> Transcendental idealism is an indispensable critical tool, not, as he repeatedly insists, a dispensable metaphysical dogma. Moreover, nowhere is this more clearly evident than in Kant’s analysis and resolution of the Antinomy of Pure Reason, which from beginning to end is methodological in nature.

Since this chapter as a whole is essentially an extended argument for this thesis, it seems appropriate to conclude it with a brief review of the main points bearing on the issue. (1) The antinomial conflict arises from neither  $P_2$  nor transcendental realism considered in isolation but from their combination. (2) It does so because transcendental realism is incapable of avoiding being deceived by this illusory principle. Indeed, it is, as it were, programmed to regard this principle as having objective validity, which, in the case of the cosmological ideas, leads inevitably to a two-sided illusion regarding “world” or “nature”

considered as self-contained totalities. (3) But transcendental realism is not itself a metaphysical position in the traditional sense; it is rather a standpoint, shared by proponents of diverse metaphysical views, from which both metaphysical and epistemological issues are approached. Specifically, it may be understood as the standpoint that approaches these issues (including the cosmological questions) from a God's-eye or theocentric perspective, which systematically ignores the role of spatiotemporal conditions in the conception of how such totalities are "given."<sup>82</sup> (4) Since transcendental realism is a metaphilosophical standpoint rather than a metaphysical doctrine, the transcendental idealism to which Kant opposes it must be granted a similar status. (5) Consequently, the indirect proof of transcendental idealism, through a demonstration of the contradictory consequences yielded by transcendental realism in its engagement with the sensible world considered as a whole existing in itself, must be seen as an argument for the necessity of adopting the opposed standpoint of transcendental idealism in order to deal successfully with the cosmological problematic, not as an argument for a phenomenalist metaphysical doctrine combined with an unknowable realm of things in themselves.<sup>83</sup> As Kant puts the matter at the end of his account of transcendental idealism as the key to the resolution of the Antinomy, the question of whether we are dealing with appearances or things in themselves does not arise within the context of an empirical investigation of nature. Rather,

[o]nly in another relation, when these same appearances are to be used on behalf of the cosmological idea of an absolute whole and having to do with a question that goes beyond the bounds of possible experience, is it important to distinguish between the ways one might take the reality of objects of sense when thinking them, so as to prevent a deceptive delusion [*trüglichen Wahne*]<sup>84</sup> that must inevitably arise if we misinterpret our own concepts of experience. [A496–97 / B524–25]

In other words, the transcendental distinction, which constitutes the heart of transcendental idealism, is a bit of metaphilosophical therapy rather than a first-order metaphysical doctrine. In the present context, its function is to prevent the misinterpretation of empirical concepts of objects as concepts of things as they are in themselves, which, in turn, leads to the attribution of objective reality to the cosmological idea and to the consequent antinomial conflicts. If emphasizing this therapeutic function and insisting on a sharp distinction between Kant's transcendental idealism and the phenomenism or subjective idealism with which it is usually equated makes this interpretation "anodyne," then so be it.

## Chapter 14 The Ideal of Pure Reason

Kant's critique of rational theology in the *Ideal* falls roughly into three main parts: an analysis of the rational credentials of the concept of an *ens realissimum*, which forms the philosophical kernel of the idea of God; a consideration of human reason's need to assume that *some* being exists with absolute necessity and to identify that being with the *ens realissimum*; and a critique of the theoretical attempts to demonstrate this dual thesis. These are the topics of the first three parts of this chapter. A fourth part is concerned with what might be termed Kant's diagnosis of the pathology of human reason in the theological domain, where the source of the problem is once again seen to lie in the virulent combination of transcendental illusion and transcendental realism. The main question addressed throughout is: How, on the one hand, can Kant affirm the rational necessity of the *concept* of the *ens realissimum*, while, on the other, insisting on the necessarily fallacious nature of all attempts to establish its existence?

# I. THOROUGHGOING DETERMINATION AND THE *ENS REALISSIMUM*

We saw in chapter 11 that Kant distinguishes the cosmological idea and the theological idea on the grounds that the former arises from the thought of the unconditioned with respect to appearances and the latter from the thought of the unconditioned with respect to things in general. We also noted that, at that point, the need to distinguish between these two distinct ways of thinking the unconditioned, not to mention the connection of the cosmological idea with the hypothetical and the theological idea with the disjunctive syllogism, remained opaque. Thus, having seen why the cosmological idea is connected with appearances considered as constituting an absolute totality, our present concern is with the necessity of this additional thought of an absolute or unconditioned totality with respect to things in general.

Unfortunately, Kant's explanation is notoriously obscure. Instead of appealing directly to the need to assume the existence of some primal, necessarily existent being to anchor, as it were, the chain of contingency, Kant introduces what appears to be a fresh and, at least at first glance, highly recondite problem of complete or thoroughgoing determination (*durchgängige Bestimmung*). Since this problem seems far removed from the concerns of traditional theology but is one with which Kant was deeply concerned in his earliest metaphysical reflections, some older commentators have dismissed Kant's discussion as a lapse into a pre-critical dogmatic metaphysics, presumably reintroduced in the *Critique* because of the usual architectonic considerations.<sup>1</sup> Moreover, since Kant turns directly from this to the more familiar topic of reason's need to assume something that exists with absolute necessity, it has also been suggested that Kant was of two minds about the rational origin of the idea of God.<sup>2</sup>

We shall see that both objections are misguided, but our present concern is exclusively with the first. The starting point of Kant's analysis is what he terms "the principle of **thoroughgoing determination**," which states: "[A]mong **all possible** predicates of things, insofar as they are compared with their opposites, one must apply to it." This is contrasted with the "principle of **determinability**," which holds that "of **every two** contradictorily opposed predicates only one can apply to it" (A571 / B579). The main point is that, whereas the latter is merely a logical principle, reducible to the principle of contradiction,<sup>3</sup> the former is metaphysical, since it considers every thing not simply in terms of a pair of contradictory predicates but "in relation to **the whole of possibility**, as the sum total of all predicates of things in general." Moreover, by presupposing the

latter as a condition *a priori*, “it represents every thing as deriving its own possibility from the share it has in the whole of possibility” (A572 / B600).

The principle of thoroughgoing determination thus holds of *things*, not merely *concepts*, or, more precisely, it holds of the very concept of a thing. It affirms that of all possible predicate pairs (a predicate and its contradictory opposite,  $av \sim a$ ,  $bv \sim b$ , and so on) one must belong to every thing as a condition of its thinghood. Although Kant does not explicitly characterize it as such, this makes it synthetic *a priori*, which indicates that it calls for some kind of deduction.<sup>4</sup> Moreover, Kant does provide one, albeit, once again, without characterizing it as such.

The deduction turns on the conditions of the complete cognition of an individual thing, which from the timeless point of view of reason is equivalent to the conditions of its real (absolute) possibility.<sup>5</sup> The underlying idea is Leibnizian in origin, though not for that reason alien to Kant’s critical philosophy.<sup>6</sup> Simply put, it is that only a complete concept in the sense defined by the principle of thoroughgoing determination is sufficient to define an individual. In other words, such a concept must contain one member of every possible predicate pair, which means that the task of defining an individual, which for Kant as for Leibniz is an infinite one, requires an exhaustive inventory or synthesis of these pairs.<sup>7</sup> From a Kantian standpoint, the necessity of such a project derives from the nature of concepts as partial representations. Since every such concept is by its nature general, nothing less than the absolute totality of predicate pairs suffices to determine an individual. Consequently, Kant gives to the principle of thoroughgoing determination the status of a rule, itself stemming from reason, for the complete use of the understanding (A573 / B601).

Although this suggests that Kant’s appeal to this principle is rooted in his conception of human cognition as discursive, it is not a topic on which he thematizes for its own sake. Rather, his intent is to use it as the basis for a regressive argument culminating in the concept of an *ens realissimum* as the unique ideal of pure reason. Consequently, it is this regressive argument, consisting of three steps, that we must now consider.

(1) *From Thoroughgoing Determination to the Sum Total (Inbegriff) of Possibility.* In introducing the principle of thoroughgoing determination, Kant indicates that it has as its transcendental presupposition the “material of **all possibility**” (A572–73 / B600–01). The point is that as a condition of engaging in the project of exhaustively defining an individual, the understanding necessarily presupposes the sum total of possible predicates as available for the task.

This sum total may thus be thought of as an all-inclusive “predicate pool” or, in Allen Wood’s term, an “ontological space,” from which one predicate from each pair is required for the complete determination of an individual.<sup>8</sup> Underlying this line of thought is the view, traceable back to Kant’s earliest metaphysical reflections, that possibility rests not merely on a formal condition (lack of internal contradiction) but also on the “material condition” that there is some content to be thought.<sup>9</sup> Thus, in the *Critique*, it is this pool of predicates that provides the “material of all possibility,” which is presupposed by the principle of thoroughgoing determination.

(2) *From the Sum Total of Possibility to the Omnitudo Realitatis.* In order to understand Kant’s line of thought at this point, it is first of all necessary to become clear about what is meant by ‘reality’. Following a tradition that reaches back to Scotus and includes Kant’s own rationalist predecessors, Kant understands by a thing’s reality not its existence or actuality but rather its positive nature, that which defines the “whatness” (*quidditas* or *essentia*) or “thinghood” (*Sachheit*) of a thing.<sup>10</sup> Although it has scholastic roots and is closely associated with doctrines of the dogmatic rationalists, this conception of reality is already at work in the *Analytic*. In fact, it is what is thought in the category of reality, which falls under the title of quality and corresponds to the logical function of affirmation.<sup>11</sup> Moreover, it allows one to speak of reality as coming in degrees, which at the phenomenal level means having an intensive magnitude, as well as of things as composed of distinct “realities” (sensible qualities).<sup>12</sup>

In the *Ideal*, however, Kant is concerned with intelligible or noumenal rather than sensible or phenomenal reality.<sup>13</sup> This is a consequence of the purely conceptual nature of the exercise, its focus on the conditions of the possibility of the thinkability of things in general, rather than of the possibility of the cognition of things as objects of experience. Nevertheless, this does not affect the basic conceptual point, which is the priority (both logical and ontological) of realities or positive predicates over negative ones.

This makes it possible to eliminate an indefinite number of predicates from those required to conceive the possibility of a thing, namely, the negative members of the predicate pairs. It is not that these cannot enter into true propositions about a thing: we can say, for example, that a certain animal lacks cunning as easily as that it possesses agility. It is rather that, as derivative, such negative predicates, which do not express realities, can be excluded from the “transcendental substratum,” which, according to Kant, “contains as it were the entire storehouse of material from which all possible predicates of things can be taken”

(A576 / B603). Moreover, this “substratum” or “storehouse,” considered as a whole, contains “[a]ll of reality” or “*omnitudo realitatis*” (A575–76 / B603–06).

In light of this conception of the *omnitudo realitatis*, Kant invites us to consider the thoroughgoing determination of an individual as a process of iterative limitation. For example, in thinking of something as a bird I implicitly exclude all those realities that are incompatible with this; in determining it as a swan I similarly exclude those not shared by birds of that species; and in further specifying it as a female I exclude still more realities, ad infinitum. Thus, even though these negations play a crucial role in the thoroughgoing determination of any finite thing, they are all parasitic on the corresponding realities and ultimately on the *omnitudo realitatis*.

Kant’s language here, as well as in other treatments of this topic, is highly metaphorical. In a passage that is the source of Wood’s chosen metaphor of “ontological space,” Kant states: “All manifoldness of things is only so many different ways of limiting the concept of the highest reality, which is their common substratum, just as all figures are possible only as different ways of limiting space” (A578 / B606). Similarly, in his lectures Kant is quoted as using the analogy of light and reality, on the one hand, and shadow and negation, on the other, to make the same point (PV 28: 1005; 34).<sup>14</sup> The same metaphor is also used in a *Reflexion* in which Kant contrasts the rationalist view of complete determination as proceeding by the introduction of darkness as limitation of universal light with his own critical view of it as proceeding by the gradual introduction of light (conceptual determination) into darkness (the sensibly given) (R5270 18: 138–39).<sup>15</sup> Finally, in one place Kant appeals to the metaphor of an inexhaustible supply of marble from which an infinite number of statues (determinate individuals) are produced by carving away a certain part of the whole (Fort 20: 302; 390).

Although the contrast drawn in the *Reflexion* indicates Kant’s critical distance from the rationalist understanding of the project of complete determination, he consistently maintains that from the point of view of a consideration of things in general, which abstracts from the sensible condition under which things are given in intuition, reason is constrained to view the material of all possibility as a pre-given totality of realities in which determination is introduced by limitation.<sup>16</sup> Consequently, from this point of view, which is that of both the Ideal and Kant’s early metaphysics, the presupposition of the *omnitudo realitatis* is in order and the rationalist principle, “*omni determinatio est negatio*,” holds.

This also sheds light on Kant’s attempt to connect the project of thorough-



going determination with the disjunctive syllogism. As Kant puts it in a passage in which he explicitly alludes to his prior linkage of the three transcendental ideas with the three forms of syllogism, "The use of reason through which it grounds its determination of all things in the transcendental ideal is . . . analogous to its procedure in disjunctive syllogisms. . . ." (A577 / B605). Given the seemingly dogmatic tone of Kant's earlier remarks on the subject, this characterization of these two uses of reason (one real or transcendental and the other logical) as merely "analogous" might seem surprising.<sup>17</sup> Nevertheless, insofar as Kant wishes to emphasize the differences as well as the similarities between them, it is perfectly appropriate. In fact, Kant begins the paragraph in which the analogy is introduced by noting the *difference* between the disjunctive procedure of reason in the logical determination of a concept and its transcendental counterpart.

In the former case, the major premise contains a logical division of the sphere of a general concept, for example, All *A* is either *B* or *C*; the minor premise restricts this sphere to one part, *A* is not *B*; and the conclusion determines the concept through this part, *A* is *C* (A576–77 / B604–05). But, Kant continues, the general concept of a reality cannot be treated as a genus that can be specified *a priori*, since it is only through experience that we can become acquainted with the determinate species of reality that could fall under it. And from this he concludes:

Thus the transcendental major premise for the thoroughgoing determination of all things is none other than the representation of the sum total of all reality, a concept that comprehends all predicates as regards their transcendental content not merely **under itself**, but **within itself**; and the thoroughgoing determination of every thing rests on the limitation of this **All** of reality, in that some of it is ascribed to the thing and the rest excluded from it, which agrees with the "either/or" of the disjunctive major premise and the determination of its object through one of the members of this division in the minor premise. [A577 / B605]

Kant seems to be suggesting in the first part of this exceedingly dense sentence that the concept of the sum total of reality, which is operative in the "transcendental major premise" of the quasi-syllogistic procedure of thoroughgoing determination, has the logical structure of an intuition rather than of a concept of the understanding, since particular realities are to be thought as included in it rather than falling under it. Moreover, this presumably explains why reason's procedure with respect to it is *merely* analogous to its more familiar syllogistic procedure.

Nevertheless, as the second part of the sentence makes clear, Kant also wishes to insist that, in spite of this difference, *there is* an analogy between the function of reason in the two procedures.<sup>18</sup> The analogy turns on the disjunctive nature of the major premise, which provides the either/or (the possible realities) on the basis of which the thoroughgoing determination of an individual proceeds. Thus, following Wood, we might think of the major premise as composed of a comparison of an individual to be determined (*A*) with all of reality (or “ontological space”) considered as divided into two parts (*F* and *G*), each consisting of a set of realities.<sup>19</sup> Given this disjunction, the determination then proceeds by eliminating in the minor premise the realities located in one portion of this space (those contained in *G*), leaving us with the remainder (those contained in *F*) as completely and uniquely determining *A*.

As indicated above, this procedure is best described as quasi-syllogistic because of the peculiar status of the transcendental major premise. But this need not lead us to conclude with Wood that “it is difficult to regard the syllogistic procedure here as anything but a rather clumsy and artificial scholastic device.”<sup>20</sup> Although it may be natural to look at it in this way, if one dismisses the underlying theory of reason of which it is a part as itself nothing more than a heavy-handed exercise of architectonic building, it may also be viewed more charitably as pointing to a deep isomorphism between a function of reason at two levels. Indeed, as we have already seen, such a functional isomorphism is a pervasive and important feature of Kant’s thought.<sup>21</sup>

(3) *From the Omnitudo Realitatis to the Ens Realissimum*. This is the final step in the regress, albeit one that Kant does not explicitly acknowledge as such, since he tends to collapse the two lines of reasoning into one. The move turns on the all-inclusiveness of the concept of the *omnitudo realitatis*, which itself emerged from the concept of the sum total of all possibility. As occupying all of ontological space, apart from which there is no material for determination, the *omnitudo realitatis* is necessarily thought as itself thoroughly determined, that is, as defining an individual. As Kant puts it, the underlying idea of the sum total of possibility “refines itself to a concept thoroughly determined *a priori*, and thereby becomes the concept of an individual object that is thoroughly determined merely through the idea, and then must be called an **ideal** of pure reason” (A574 / B602).

It is at this point that the distinction between ideas and ideals comes into play. Previously, Kant saw no need to differentiate these two products of pure reason, but now he does so in terms of their specificity. Whereas an idea, some-

what like a concept of the understanding, refers to a kind of thing or quality, albeit qua maximized or thought in its complete perfection, for example, the idea of perfect humanity, an ideal, somewhat like an intuition, refers to an individual, albeit one constituted by purely rational predicates. It is, in Kant's terms, "the idea not merely *in concreto* but *in individuo*, i.e., as an individual thing which is determinable, or even determined, through the idea alone" (A568 / B596).

Although Kant speaks loosely of human reason as having several such ideals, of which the Stoic sage is an example, and suggests that these have normative force as exemplars, he later qualifies this by claiming that the *ens realissimum* is "the one single genuine ideal of which human reason is capable, because only in this case is an—in itself universal—concept of one thing thoroughly determined through itself, and cognized as the representation of an individual" (A576 / B604). Since we have seen that Kant regards the world both as an idea and as an individual, this may seem somewhat surprising. But it should be kept in mind that the idea of the world is not completely determined. Indeed, that is precisely the source of the antinomial conflict.

Nevertheless, it has been objected that Kant cannot regard the *ens realissimum* as a genuine ideal either. The problem allegedly lies in its incompatibility with Kant's anti-Leibnizian claim in the Amphiboly that, quite apart from any logical contradiction, realities might conflict with each other, that is, exhibit a "real opposition" (A272–73 / B328–29), which would preclude the very possibility of the *ens realissimum*.<sup>22</sup> Kant could hardly have been unaware of this problem, however, and he presumably attempted to deal with it by modifying what he terms the "first crude outline" of his account (A579 / B607). In the initial version, the all of reality was regarded as a totality in which particular realities are contained as discrete portions of ontological space, which naturally raises the question of their mutual compatibility. In the refined version of his account, however, which is intended to link the concept of an *ens realissimum* to the God of rational theology, Kant states:

[T]he highest reality would ground the possibility of all things as a **ground** and not as a **sum** total; and the manifoldness of the former rests not on the limitation of the original being itself, but on its complete consequences; to which our whole sensibility, including all reality in appearance, would then belong, which cannot belong to the idea of a highest being as an ingredient. [A579 / B607]

Although the latter part of this passage indicates that Kant's prime concern was to avoid the Spinozistic implications of the identification of God with the

sum total of reality, the substitution of the grounding for the totalizing metaphor addresses our present concern as well. That is to say, if the *ens realissimum* is conceived as the ground of all reality rather than as the sum total thereof, it may be possible to preserve the concept from the threat of containing incompatible predicates. To be sure, this does not suffice either to establish the real possibility of the *ens realissimum* or to make comprehensible the nature of the relation between such a ground and the realities that are supposedly grounded therein. Nevertheless it does, at least, preserve the conceivability of such a being, which is all that Kant is concerned to do at this point.

This much is clear from Kant's initial distancing of his analysis of the generation of the ideal from any claims about existence. As he puts it immediately upon completing his analysis of the conditions of the possibility of a thoroughgoing determination, which supposedly culminates in the concept of the *ens realissimum*:

It is self-evident that with this aim—namely, solely that of representing the necessary thoroughgoing determination of things—reason does not presuppose the existence of a being conforming to the ideal, but only the idea of such a being, in order to derive from an unconditioned totality of thoroughgoing determination the conditioned totality, i.e., that of the limited. [A577–78 / B605–06]<sup>23</sup>

In other words, the argument up to this point is to be taken as an exercise in conceptual analysis, designed to show that the concept of the *ens realissimum* is a necessary presupposition of reason, rather than as a putative demonstration of the existence of a being corresponding to this concept. Nevertheless, this does not prevent the concept from being a product of transcendental illusion. Indeed, it must be viewed as such because of its genesis. This becomes evident as soon as one recognizes that the move from the sum total of possibility to the *omnitude realitatis*, and from this to the *ens realissimum*, turns on the ubiquitous and unavoidable commitment to  $P_2$ . Here, as elsewhere, it is  $P_1$  that underlies the search for conditions, this time of the complete determination of an individual, and once again this search is inseparable from the assumption that the complete totality of conditions is “given” ( $P_2$ ). But to assume this is precisely to assume the givenness of the *omnitude realitatis*, which, as constituting a completely determined individual, is, in turn, equivalent to the *ens realissimum*. Thus, the latter, as a subjectively necessary product of pure reason, is also a product of transcendental illusion.

As we have also seen, however, illusion is not yet error, which is just the point that Kant is making with his sharp distinction between reason's formation of

the *idea* of the *ens realissimum* (the Ideal) and the assumption of its existence. It is the latter, which Kant terms ‘hypostatization’, that constitutes the basic metaphysical error. Consequently, Kant’s concern is to analyze the process leading to this hypostatization and to the subsequent identification of the *ens realissimum* with a being whose existence is absolutely necessary. It is to this analysis, which constitutes the heart of the Ideal, that we now turn.

## II. THE HYPOSTATIZATION OF THE IDEAL AND ABSOLUTELY NECESSARY EXISTENCE

As Kant describes his diagnostic task,

It is not enough to describe the procedure of our reason and its dialectic; one must also seek to discover its sources, so as to be able to explain this illusion itself, as a phenomenon of the understanding; for the ideal we are talking about is grounded in a natural and not merely arbitrary idea. Therefore, I ask: How does reason come to regard all the possibility of things as derived from a single possibility, namely that of the highest reality, and even to presuppose these possibilities as contained in a particular original being? [A581 / B609]

Like much else in the *Critique*, Kant’s answer to these questions is embedded in a complex, yet extremely cryptic, analysis, which lends itself to a variety of interpretations. This analysis is contained in the remainder of the second and the third sections of the Ideal and falls into two parts: the first dealing with the issue of hypostatization and the second with the attribution to the *ens realissimum* of absolutely necessary existence. We shall see that these are two chapters in a single story: the first treating the how or what we might term the mechanism of hypostatization and the second the why or the motivation for it.<sup>24</sup> To anticipate, in spite of its subjective necessity qua idea, if reason did not need to posit something that exists *necessarily*, it would have no compelling reason to attribute existence to the *ens realissimum* at all.

### A. The Mechanism of Hypostatization

The initial question is how existence comes to be attributed to the *ens realissimum*, that is, how the idea is hypostatized. At this stage of the analysis, Kant’s account closely parallels that of the Paralogisms, which, it will be recalled, Kant described as based on the “subreption of hypostatized consciousness” (A402). There it was the I of apperception, whereas now it is the *omnitudo realitatis*, under the guise of the *ens realissimum*, that is hypostatized. Nevertheless, in both

cases, this involves an illicit use of principles whose home turf is supposedly the Transcendental Analytic.

In the present case, Kant begins by asking rhetorically: Why, since the necessary determining function of the *ens realissimum* presupposes merely the idea of such a being and, therefore, could be performed by a “mere fiction” (A580 / B608), is a real existence attributed to it? And he responds, “The answer suggests itself on the basis of the discussions of the Transcendental Analytic themselves” (A581 / B609).

Although Kant does not specify what discussions he has in mind, his point seems to be that the possibility of experience, like that of things in general, rests on material as well as formal conditions. Consequently, even though the overall concern of the Analytic was with the latter, further reflection shows that its argument presupposes that the material of empirical possibility, namely, “reality in appearance (corresponding to sensation),” is given. Moreover, since, as is emphasized in the Analytic, “[t]here is only one experience . . . just as there is only one space and time” (A110), all this material must be given in “the one all-encompassing experience.” And from this it supposedly follows that “the material for the possibility of all objects of sense has to be presupposed as given in one sum total; and all possibility of empirical objects . . . can rest only on the limitation of this sum total” (A582 / B610).

This amounts to an argument for the thesis that an analogue of the *omnitudo realitatis* must be presupposed at the empirical level. Moreover, given the assumption that possibility rests upon material conditions, which is traceable back to Kant’s earliest metaphysical speculations, together with the critical conception of the unity of experience, this claim seems in order. Indeed, it even entitles Kant to affirm the principle that “nothing is an object **for us** unless it presupposes the sum total of empirical reality as condition of its possibility” (A582 / B610). As Kant continues, however, the problem is that,

[i]n accordance with [*nach*] a natural illusion, we regard as a principle that must hold of all things in general that which properly holds only of those which are given as objects of the senses. Consequently, through the omission of this limitation we will take the empirical principle of the possibility of our concepts of things as appearances to be a transcendental principle of the possibility of things in general. [A582 / B660]

If the previous analysis is correct, the “natural illusion” to which Kant alludes consists in the transformation of the sum total of empirical possibility, which is necessarily presupposed as a condition of thoroughgoing determination, into the transcendental idea of the *ens realissimum*. The conception of the *omnitudo*

*realitatis* is an intermediate stage in this dialectical process, whose function is to provide, as it were, the content for this idea. Consequently, Kant's claim is that this illusion somehow leads to (without being identical with) the metaphysical error, which once again consists in an illicit hypostatization.

This still leaves us, however, with a set of questions similar to those we considered in chapter 12 in connection with the Paralogisms. Specifically, we need to see, first, how the illusion leads to the conflation and, second, how the latter is connected with the hypostatization. Since the conflation of an empirical with a transcendental principle is presumably avoidable, it must be distinguished from the underlying illusion; but since it does not, of itself, involve an existence claim, it must also be distinguished from the hypostatization. The closest that Kant comes to spelling out the details of this complex story is in the final paragraph of this section, where he writes:

That we subsequently hypostatize this idea of the sum total of all reality, however, comes about because we dialectically transform the distributive unity of the use of the understanding in experience, into the collective unity of a whole of experience; and from this whole of appearance we think up an individual thing containing in itself all empirical reality, which then—by means of the transcendental subreption we have already thought—is confused with the concept of a thing that stands at the summit of the possibility of all things, providing the real conditions of their thoroughgoing determination. [A582–83 / B610–11].

Here Kant suggests that the dialectical process leading to the hypostatization of the idea of the *omnitudo realitatis*, under the description of the *ens realissimum*, consists of three logically distinct steps: (1) the transformation of a distributive into a collective unity; (2) the cogitation of an individual thing that somehow contains all of empirical reality; and (3) the confusion, by means of the previously considered transcendental subreption (the conflation), of this cogitated high order empirical individual with the *ens realissimum*. Presumably, the third step constitutes the actual hypostatization, and the first two steps its necessary preconditions.

In order to understand the first step, we must keep in mind that the distributive unity corresponds to the unity of the understanding and the collective unity to that of reason.<sup>25</sup> The former unity is distributive in the sense that it is produced by the unification of discrete phenomena, “distributed” in space and time, under a universal rule by virtue of which they become ingredients in the “one all-encompassing experience” (A582 / B610). Thus, the fact that the material of empirical possibility constitutes a distributive unity is a direct conse-

quence of the singleness of the experience in which this material is located. By contrast, the unity of reason is collective in the sense that it “collects” these rules, and indirectly the phenomena standing under them, into a systematic whole connected by rules of inference. Accordingly, to transform the former into the latter “dialectically” is to convert a principle holding at the empirical level into one applicable to things in general. In Kant’s terms this amounts to a subreption.

So described, this subreptive conversion is equivalent to the conflation referred to above. Consequently, the first step basically redescribes the result already attained rather than adding anything materially new. Under either description, however, the systematic importance of this step consists in the fact that it is here that transcendental realism first enters the picture. Admittedly, Kant does not inform us of this fact (things might have been somewhat clearer if he had); nevertheless, given our previous analysis of such realism, it seems to be a reasonable inference and points to a significant parallelism with the account of the pathology of the rational psychologist sketched in the Paralogisms.

What links this conflation to transcendental realism is that it results from the failure to distinguish between an empirical principle concerning things as objects of possible experience and a transcendental principle concerning things in general. Nor should one be led astray by the fact that Kant here refers to a conflation between a principle concerning things as they appear and things in general, whereas transcendental realism has been seen to consist in the identification of appearances with things in themselves. Since to consider things in general just is to consider them in abstraction from the conditions under which they appear, this is equivalent to considering them as they are in themselves. Thus, in making claims about the conditions of things in general, the rational theologian is proceeding as a transcendental realist. Moreover, proceeding in this way, the transcendental realist inevitably (and from such a realist’s point of view, quite reasonably) regards conditions of appearances as if they were conditions of things in general. The conflation with which Kant is concerned here is clearly a product of that mistake. Although not itself a substantive metaphysical error, it might be described as a “methodological error,” which, under the influence of the illusion, leads inevitably to the latter.

The second step, in which the concept of an individual thing containing the whole of appearance is introduced, may be regarded as the marriage of this conflation with the underlying illusion. As such, it provides the necessary precondition of the hypostatization. The illusion is involved because the idea of the whole of appearance as constituting a higher-order thing or individual is itself



illusory. In fact, it is the very illusion that was operative in the generation of the cosmological idea of the world.<sup>26</sup> Unlike the Antinomy, however, where the spatiotemporal conditions of experience were factored in and the idea was, therefore, pseudo-empirical, here the idea may be described as “pseudo-rational.”<sup>27</sup> Consequently, there is no antinomial conflict here. Instead, the transcendental realist’s conflation (made under the spell of the transcendental illusion) leads to the idea of an individual as the ground of an empirical reality that is illicitly identified with the reality of things in general.

Perhaps the most perplexing feature of this deeply obscure account is the apparent unnaturalness of the thought that the sum total of empirical reality constitutes an individual in any meaningful sense, much less one that might be conflated with the subjectively necessary idea of the *ens realissimum*. It becomes somewhat less mysterious, however, if we take Kant to be referring to the dialectically transformed sum total of empirical reality rather than the one initially presupposed as a condition of experience. As dialectically transformed (through the transcendental realist’s conflation), this sum total becomes the *omnitude realitatis*. But as such it is an idea of reason, and it is only by being taken in this way that it can be considered to be thoroughgoingly determined.

Finally, the actual hypostatization of an *ens realissimum*, which occurs in the third step, may be seen as the issue of this marriage of transcendental illusion and transcendental realism. Simply put, the former provides the idea to be hypostatized (that of the *ens realissimum*) and the latter the essential precondition of its hypostatization.

In a note attached to the paragraph cited above, Kant clarifies matters somewhat by introducing a distinction between realization and hypostatization.<sup>28</sup> As he puts it there, “This ideal of the supremely real being, even though it is a mere representation, is first realized, i.e., made into an object, then hypostatized, and finally . . . it is even personified” (A583 / B611n.). Although personification can be ignored here, the contrast between the realization and hypostatization is of paramount significance, since it amounts to the distinction between the unavoidable result of the illusion and the avoidable result of being deceived by it. Of particular relevance is the characterization of realization as an act of making an idea into an object. Since to hypostatize is to give an object a real existence independent of its idea, insofar as realization is distinguished from it, the objectification it involves can be understood only as the generation of a merely intentional object or what Kant terms an “object [given] in the idea,” as opposed to one given “absolutely” (A670 / B698).<sup>29</sup> Consequently, the hypostatization, which occurs on the basis of transcendentially realistic assumptions,

may also be viewed as attributing absolute, that is, real, existence to something that exists only in idea. In the case of the Paralogisms it was the idea of the soul, and here it is that of an *ens realissimum*.

### **B. Absolutely Necessary Existence and the *Ens Realissimum***

The obvious problem posed by the preceding analysis is that this hypostatization seems so transparently fallacious that, unlike the case with some of the reasoning in the Paralogisms, it is difficult to see how anyone could succumb to it. But rather than viewing this as an objection to his account, it is a point on which Kant himself insists in the third section of the Ideal. As he puts it at the beginning of this section,

In spite of its urgent need to presuppose something that the understanding could take as the complete ground for the thoroughgoing determination of its concepts, reason notices the ideal and merely fictive character of such a presupposition much too easily to allow itself to be persuaded by this alone straightway to assume a mere creature of its own thinking to be an actual being, were it not urged from another source to seek somewhere for a resting place in the regress from the conditioned, which is given, to the unconditioned, which in itself and as regards its mere concept is not indeed actually given, but which alone can complete [the] series of conditions carried out to their grounds. [A583–84 / B611–12]

This suggests that the function of this section is to uncover this “other source,” which is precisely what Kant attempts to do. Fortunately, in contrast to the recondite analysis of section 2, section 3 covers territory that is much closer to the considerations that are usually thought to lead to the positing of a supreme being. In fact, Kant now takes himself to be charting “the natural course of human reason, even the most common,” though he hastens to add that “not everyone perseveres in it” (A584 / B612). This course begins with common experience and thus takes itself as grounded in something existing rather than mere concepts. But, as Kant metaphorically puts it, “this ground [*Boden*] gives way unless it rests on the immovable rock of the absolutely necessary” (A584 / B612). In other words, what naturally leads human reason to the thought of the *ens realissimum* is not the philosophical project of locating the transcendental conditions of the possibility of thoroughgoing determination, but merely the need to assume something that exists with absolute necessity in order to find a resting place for thought. The operative term here is ‘absolute’, since reason’s concern is not with something whose existence follows necessar-

ily from an antecedent cause (that is the business of the understanding), but rather with something whose existence is unconditionally necessary, that is, necessary in all respects or in virtue of its inherent nature.<sup>30</sup>

According to Kant, underlying this project of common human reason is the twofold illusion that such a resting place is attainable at all and that it is uniquely provided by the *ens realissimum*. Section 3 is devoted to exposing the latter, and the subsequent two sections, dealing with the ontological and cosmological proofs, to the former. For the present we shall focus on the latter.

Kant begins by stipulating what he takes to be the fundamental assumption of common human reason, which he initially treats as if it were non-problematic. Simply put, it is that if one admits that anything at all exists, then one must also concede that something exists with absolute necessity. This is because reason is constrained to think something that exists in this manner as the ultimate ground of the contingent (A584 / B612). At one point, Kant suggests that this reasoning, which arises as soon as one reflects on the causes of things, is based on the thought of the “inner insufficiency of the contingent” (A589 / B617). It is also the basis of the familiar cosmological proof, which Kant presents as the first and most natural from the standpoint of ordinary human reason.<sup>31</sup>

Given this assumption, the task is to find a candidate on which to bestow the exalted status of absolutely necessary existence. By its very nature, this is a purely conceptual endeavor, and it is at this point that the concept of the *ens realissimum* enters the picture. Once again, the argument is simple and straightforward, as befits common human reason in its initial venture into the transcendent. Since, *ex hypothesi*, the *ens realissimum* contains all reality, there can be nothing on which it depends. Accordingly, it leaves no room for any further “why” questions, which is precisely what is required of something that supposedly exists with absolute necessity.

Although Kant clearly rejects this line of argument, he also expresses considerable sympathy with it. In fact, he attributes to it “a certain importance and high regard of which it cannot straightway be divested simply on account of this objective insufficiency” (A588–89 / B616–17). Moreover, this sympathy has a deep philosophical basis in the analysis of the *ens realissimum* as the “one single genuine ideal of which human reason is capable” (A576 / B604) and, indeed, as a “faultless ideal, a concept which concludes and crowns the whole of human cognition” (A641 / B669). Thus, in view of this analysis, Kant admits that

[i]t cannot be disputed that this concept has a certain cogency if it is a matter of making **decisions**, that is, if the existence of some necessary being is already conceded, and one agrees that one must take sides on where one is to place it; for then one can make no more suitable choice than—or rather, one has no other choice, but is compelled—to vote for the absolute unity of complete reality as the original source of possibility. [A587 / B615]

The point, however, is that no such decision is required, since the antecedent is not to be conceded. Nevertheless, rather than introducing that large issue at this juncture, Kant focuses on common reason's assumption that the *ens realissimum* is the only *conceivable* candidate for absolutely necessary existence. In spite of the pristine nature and subjective indispensability of the ideal, Kant believes that this assumption is open to question on the grounds that there is no contradiction in attributing absolute necessity to a limited being not possessing all of reality. "For even if in its concept I do not find the unconditioned which the All of conditions already carries with it, still it cannot be concluded from this that its existence must be conditioned" (A588 / B616).

Kant's point is not that it is plausible to attribute absolutely necessary existence to a limited being, merely that there is no contradiction in doing so. He illustrates this by appealing to the analogy with the hypothetical syllogism, where the negation of the antecedent in the minor premise does not legitimize the negation of the consequent in the conclusion. In other words, even if (for the sake of argument) it is assumed that being the *ens realissimum* is a *sufficient* condition for the attribution of absolutely necessary existence, it does not follow that it is also a *necessary* condition. Thus, the inference to the *ens realissimum* as the ultimate ground of conditioned existence is seen to rest on a non sequitur. Moreover, this effectively removes one of the two major props of rational theology. The removal of the second prop (the concept of an absolutely necessary existence) is the task of the critiques of the ontological and cosmological proofs.

### III. THE ONTOLOGICAL AND COSMOLOGICAL PROOFS

Schopenhauer famously claimed that Kant did not have to write the *Critique of Pure Reason* in order to refute the ontological argument.<sup>32</sup> And he was certainly correct, since not only many previous philosophers but also Kant himself, in his earliest metaphysical writings, had already refuted it in its classical Anselmian-Cartesian form.<sup>33</sup> Nevertheless, Kant's critique remains of interest because, on

the one hand, through its denial that 'existence' is a real predicate, it has provided the basis for many subsequent critiques, while, on the other, it is closely intertwined with his systematic account of reason and its unavoidable illusions. Indeed, what is distinctive about Kant's critique is precisely that he takes seriously the problematic notion of absolutely necessary existence. Rather than an absurdity to be dismissed summarily, as it was for Hume, this notion is for Kant the "true abyss" for human reason (A613 / B641). But since the latter point is fully brought out only in the critique of the cosmological proof, we must consider this critique as well.<sup>34</sup>

### A. The Ontological Proof

Kant's critique, which appears to be aimed at both the Cartesian and Leibnizian versions of the proof, falls neatly into two parts: general considerations that speak against the attribution of absolutely necessary existence to *any* entity, and a refutation of the claim that the concept of the *ens realissimum* is an exception to this rule. The entire discussion is prefaced, however, by the reflection, based on the preceding analysis, that "the inference from a given existence in general to some absolutely necessary being seems to be both urgent and correct"; the problem being that "in framing a concept of such a necessity, we have all the conditions of the understanding entirely against us" (A592 / B620). In other words, we seem to be confronted with another instance of the conflict between the claims of reason, for which the concept of an absolutely necessary being is indispensable (though illusory), and the conditions of the understanding, which has no room for it.

The general considerations to which Kant appeals here, and which he treats as non-problematic, concern mainly the nature of existential judgment and its various modalities. Although he does not formulate it as such, the account is based on this principle: For every predicate, if the subject is rejected with the predicate, there is no contradiction. Since by an absolutely necessary being is meant one whose existence cannot be denied without contradiction, it follows that no being can be judged necessary in that sense.

The reason there is no contradiction if the subject of a proposition is denied together with its predicates is simply that there is then nothing left to contradict. If *x* is a triangle, it follows necessarily that it has three angles. One cannot affirm the subject of a judgment and deny without contradiction the predicates that necessarily pertain to it. But if there is no subject of which these angles are predicated, there is no contradiction.

As already noted, Kant assumes that most philosophers would readily ac-

knowledge this as a general principle, but that some might wish to insist on an exception in the case of the concept of the *ens realissimum*. Thus, the goal of the second part of the critique is to deny that it is an exception. The argument in favor of an exceptional status for the *ens realissimum* is familiar, and so is Kant's criticism of it. Simply put, the pro argument turns on the premise that existence is itself a reality or perfection. Consequently, assuming that such a being is possible, which Kant is willing to grant for the sake of the argument,<sup>35</sup> it follows that such a being must exist with absolute necessity, because any *ens realissimum* that did not exist in this way would, *ex hypothesi*, not be the *ens realissimum*. Moreover, this argument applies uniquely to such a being, since there would be no contradiction in denying the existence of any being possessing less than all of reality.

Given this line of argument, the obvious retort is to deny that 'existence' is a predicate or, more precisely a "real predicate" or "determination," as opposed to a merely logical predicate, which is exactly what Kant does. The very possibility of existential judgments shows that 'existence' is at least a logical predicate, but it does not suffice to show that it is anything more than that.

Although Kant's denial that 'existence' is a real predicate is widely accepted among philosophers, including many who find little else in the *Critique* to their tastes, it has not gone unchallenged. The problem is a very broad and complex one, however, which cannot be dealt with here. Nevertheless, it is important to consider briefly two objections that are directed specifically to Kant's analysis rather than the broader issue.

One objection that is easy to dismiss concerns an apparent incompatibility between Kant's denial that 'existence' is a real predicate, on the one hand, and his insistence that all existential judgments are synthetic and that synthetic judgments involve determinations or real predicates, on the other. In short, the objection is that these constitute an inconsistent triad. If existential judgments are synthetic, then, by Kant's own criterion of syntheticity, 'existence' must function in such judgments as a real predicate. Conversely, if it does not, then it must follow on Kantian principles that such judgments are analytic, which is absurd.<sup>36</sup>

The problem has already been touched upon in chapter 4. As was noted there, what makes an existential judgment synthetic is not it that affirms 'existence' as a real predicate but rather that, like all synthetic judgments, it makes a claim about the subject that goes beyond what is contained in its concept, namely, that it exists or that the concept is instantiated. Consequently, all that is required at this point is a clarification of Kant's account of syntheticity. As be-

fore, what makes any judgment synthetic is that it “materially” extends our knowledge beyond what is already thought (implicitly or explicitly) in the concept of its subject. But it can do this in either of two distinct ways: by affirming (or denying) a further determination of the subject or by affirming (or denying) that the concept of the subject is instantiated. Existential judgments are synthetic in the latter way.

A somewhat more serious objection concerns the argument Kant adduces for the denial that ‘existence’ is a real predicate, which is supposedly contained in the following passage:

Thus when I think a thing, through whichever and however many predicates I like (even in its thoroughgoing determination), not the least bit gets added to the thing when I posit in addition that this thing is. For otherwise what would exist would not be the same as what I have thought in my concept, but more than that, and I could not say that the very object of my concept exists. Even if I think in a thing every reality except one, then the missing reality does not get added when I say the thing exists, but it exists encumbered with just the same defect as I have thought in it; otherwise something other than what I thought would exist. [A600 / B628]

The objection to this line of argument, which was initially advanced by Jerome Shaffer and later endorsed (in a somewhat different form) by Allen Wood, is that if it proves that ‘existence’ is not a real predicate, it also proves that nothing could be one. This is because even real predication in a sense “changes the subject,” but this does not lead us to conclude that it is no longer the same thing. For example, if I say that some thing is red or omnipotent, then I am adding this predicate to my concept of the thing, but it remains the very same thing which I formerly thought of without that quality. But if this holds of manifestly real predicates, such as ‘redness’ or ‘omnipotence’, why should it not also hold of ‘existence’?<sup>37</sup>

The answer is simple, though it involves turning the critic’s question on its head. First, to predicate ‘existence’ of a thing presupposes its thoroughgoing determination because the latter is a condition of its thinghood. This is not to suggest that one’s *concept* of the thing judged to exist must itself be thoroughly determined, since this is possible only in the case of the concept of the *ens realissimum*, which is so determined *a priori*. The point is rather that the *thing itself* must be assumed to be so determined; otherwise the question of its existence could not meaningfully arise. But if ‘existence’ were a real predicate, analogous to ‘redness’ or ‘omniscience’, such that in the process of determining some *x* one were to assume that *x* is either red or not red, omniscient or non-

omniscient, existent or nonexistent, and so on, then it would follow that for any  $x$ , if it exists at all, it exists necessarily. Nor can this conclusion be evaded by claiming that this  $x$  no more exists necessarily than it is necessarily red. For on this picture *it is* necessarily red, since being red partially defines what it is to be that particular thing. Consequently, the Kantian retort to the objection sketched above is that one must either deny that 'existence' is a real predicate or admit that everything that exists, does so necessarily.

In response, the defender of the ontological argument will doubtless still insist that this holds only of the *ens realissimum*, since only such a being possesses *all* perfections. Any less than perfect being, even one lacking only a single perfection, might lack the perfection of existence and, therefore, not exist. Thus the contingency of existence in all cases other than that of the *ens realissimum* is preserved. But this will not work, for the very reasons Kant advances in the passage under consideration. The problem is that in the case of the *ens realissimum*, its concept, not merely the thing itself, must be assumed to be thoroughlygoingly determined *prior* to raising the question of its existence. Indeed, this is the basis for the privileged status of the concept assumed by the defender of the ontological argument. In that case, however, by including existence among these perfections one is either uttering a tautology (a being that includes existence among its properties includes existence among its properties) or changing the initial concept. Either way, the claim for a unique status for the *ens realissimum* with respect to existence cannot be maintained, which is just Kant's point.

This result also points back to Kant's analysis of modality, first discussed in chapter 6. As we saw there, Kant insists that "[t]he modality of judgments is a quite special function of them . . . in that it contributes nothing to the content . . . of the judgment, but rather concerns only the value of the copula in relation to thinking in general" (A75 / B99–100). In other words, the modal functions presuppose a fully formed judgment and assess its epistemic import. Similarly, in the Postulates, Kant maintains that "[t]he categories of modality have this peculiarity: as a determination of the object they do not augment the concept to which they are ascribed in the least, but rather express only the relation to the faculty of cognition" (A219 / B266). From this point of view, to treat 'existence' as a real predicate is to commit a category mistake, consisting in the conflation of a category of quality (reality) with one of modality (existence or actuality). This is not, of course, to claim either that Kant is correct in his denial that 'existence' is a real predicate or that this denial cannot also be made on non-Kantian grounds. It is, however, to suggest that Kant's critique of the on-



tological argument is deeply rooted in some of the central tenets of his epistemology and cannot be neatly separated from this context.

## B. The Cosmological Proof

What Kant terms the “cosmological proof” is basically a reprise of the argument sketched in section 3, where it was attributed to the “natural course of human reason.” Now, however, it appears in its proper philosophical dress as a full-fledged proof designed to avoid the main problem of the ontological proof by anchoring the argument securely in experience.<sup>38</sup> Consequently, Kant’s critique is an attempt to show that this anchoring is illusory in several respects.

As before, the proof consists of two parts: (1) an inference from the factual premise that something exists (oneself) to the conclusion that something else must exist with absolute necessity; (2) the identification of this absolutely necessary being with the *ens realissimum*. Although the point is often lost sight of, here Kant is concerned with both parts of this argument, in contrast to section 3, where his focus was almost entirely on the second part. Moreover, his treatment of the second part differs significantly in the two places. As already indicated, whereas in the earlier discussion Kant challenged the claim of the *ens realissimum* to be a necessary condition for the attribution of absolutely necessary existence, now he argues in effect that it cannot be a sufficient condition either, and that the problem lies in the concept of a being whose existence is absolutely necessary.

The best known and most widely discussed feature of Kant’s critique of the cosmological proof, indeed, the one he himself emphasizes, is his assertion of its inseparability from the ontological proof to which it is intended as an alternative. It is not that Kant simply reduces the former to the latter, since it clearly has a quite distinct starting point; it is rather that at a crucial juncture the cosmological must rely on the ontological proof. Consequently, it can neither claim to be logically independent of the ontological proof nor pretend to have a genuinely empirical basis. In fact, Kant characterizes the latter claim as a mere ruse, since the appeal to experience occurs only at the first step and plays no role in determining the nature of that which is assumed to exist with absolute necessity (A606 / B635).<sup>39</sup>

The partial convergence of the ontological and cosmological proofs stems from their common appeal to the concept of the *ens realissimum*. The fact that the one reasons *from* and the other *to* this concept makes no essential difference for Kant, since they both end up relying on the assumption that absolutely nec-

essary existence pertains uniquely to the object of this concept. Kant makes the point first informally and then in a “scholastically correct way.” Expressed in the latter way, its “*nervus probandi*” is the proposition: “Every absolutely necessary being is at the same time the most real being” (A608 / B636). Kant notes that, as a universal affirmative proposition, it is convertible *per accidens* to “Some most real beings are at the same time absolutely necessary beings.” But since, *ex hypothesi*, a most real being contains all reality, it follows (by the identity of indiscernibles) that there can be only one such being. Consequently, it is convertible absolutely to the universal proposition: “Every most real being is a necessary being.” As Kant proceeds to point out, however, this is just what the ontological proof maintains (A608 / B636). Thus, if that fails, the cosmological proof fails as well.

Nevertheless, this is not all that Kant finds wrong with the cosmological proof. In fact, he suggests that the argument contains “an entire nest of dialectical presumptions . . . which transcendental criticism can easily discover and destroy,” of which he proceeds to list four:

- 1) The transcendental principle of inferring from the contingent to a cause, which has significance only in the world of sense, but which outside it does not even have a sense. . . .
- 2) The inference from the impossibility of an infinite series of causes given one upon another to a first cause, which the principles of the use of reason cannot justify inferring within experience, still less our extending this principle to somewhere beyond it (into which the causal chain cannot be extended at all).
- 3) The false self-satisfaction reason finds in regard to the completion of this series by the fact that one finally does away with every condition—without which there can be concept of necessity. . . .
- 4) The confusion of the logical possibility of a concept of all reality united (without internal contradiction) with its transcendental possibility, which requires a principle of the feasibility of such a synthesis, but which once again can apply only to the field of possible experiences, etc. [A609–10 / B637–38]

The addition of “etc.” indicates that Kant took himself to have given merely a sample of a potentially much longer litany of philosophical sins committed by the cosmological proof. Clearly, one could discuss each of these objections at length and perhaps defend certain versions of the cosmological proof against at least some of them. The significance of Kant’s account does not, however, lie in its details, which he leaves to his readers to work out for themselves, but rather in its broad characterization of the cosmological proof as conflicting at numerous points with the conditions of the understanding. In each case, it is a matter of reason, in its endemic quest for a being that exists with absolute necessity, coming into conflict with these conditions. Moreover, all of this is independent

of the objection that the cosmological proof is forced to rely on the ontological in its second part.

Nevertheless, from the point of view of a Kantian diagnosis of the pathologies of reason, this actually heightens the interest of the proof, which lies in its revelatory value rather than its probative force. As already noted, what it reveals is the “true abyss” of human reason. Kant calls attention to this in the following well-known passage:

The unconditioned necessity, which we need so indispensably as the ultimate sustainer of all things, is for human reason the true abyss. Even eternity—however awful the sublimity with which a Haller might portray it—does not make such a dizzying impression on the mind; for eternity only **measures** the duration of things, but it does not sustain that duration. One cannot resist the thought of it, but one also cannot bear it that a being that we represent to ourselves as the highest among all possible beings might as it were, say to itself: “If I am from eternity to eternity, outside me is nothing except what is something merely through my will; **but whence** then am I?” Here everything gives way beneath us, and the greatest perfection as well the smallest, hovers without support before speculative reason, for which it would cost nothing to let the one as much as the other disappear without the least obstacle. [A613 / B641]

As this passage suggests, it is the cosmological, not the ontological, proof that directly addresses human reason’s deep need to assume something that ultimately sustains the being and duration of things. This is precisely why this proof expresses the “natural course of human reason.” It also suggests that nothing can satisfy this need, not even the *ens realissimum*, who is led on behalf of the same human reason to ask: “[w]hence then am I?” Since only a being whose existence is absolutely necessary can serve as the required ultimate sustainer, the abyss of human reason is that it both requires and cannot recognize anything that exists in that manner. At this point, however, there appear to be only two options: either the “sacrifice of the intellect” to some “higher truth” or a critical examination of the causes of this abyss and the determination of a means of escape. Fortunately, Kant chose the latter.

#### IV. Critical Analysis and Cure

Kant’s analysis of the causes of human reason’s predicament underlying the failed theistic proofs and his proposed remedy are contained in the brief section entitled “Discovery and explanation of the dialectical illusion in all transcendental proofs of the existence of a necessary being” (A614–20 / B642–48). As

one might suspect, the prime source of this predicament is the underlying transcendental illusion. Not only does such illusion lead to the previously discussed hypostatization of the ideal, it also motivates reason's quest for an absolutely necessary being, which is the root of the present difficulty.

In light of what we have already learned, the basic point can be put quite simply, and Kant does not find it necessary to dwell on it at length. It is that the ubiquitous  $P_2$  both requires reason to assume that *some* being exists in that manner and prohibits it from assuming that *any particular* being, including an *ens realissimum*, does so.  $P_2$  requires this insofar as it demands something unconditioned with respect to contingent existence; it precludes it insofar as it also requires that the question of *its condition* be raised with respect to every candidate for unconditionality.<sup>40</sup> The fallacies committed in the ontological and cosmological proofs are the result of being deceived by this illusion.<sup>41</sup>

Although Kant does not make it explicit, it is also clear that the latter is once again the result of the rational theologian's underlying transcendental realism, which here, as always, makes it impossible to avoid being deceived by the illusion. For example, the "nest of dialectical assumptions" to which Kant refers in his critique of the cosmological proof is obviously rooted in transcendental realism, since it involves an illicit extension of the categories, which is itself the consequence of a failure to distinguish between conditions of things as appearances and conditions of things in general. More generally, it is such realism that fuels the conflation of the subjective necessity (stemming from  $P_2$ ) of thinking of an unconditionally necessary ground of things in general with the demonstration of the objective necessity of the existence of such a ground. Accordingly, in order to avoid this illusion and the fallacies it generates, it is necessary to jettison this realism, which is equivalent to adopting a transcendently idealistic stance.

Against this it might be objected that the cure Kant actually proposes makes no reference to either form of transcendentalism. Instead, he suggests that the cure consists in the recognition that "[i]f I must think something necessary for existing things in general but am not warranted in thinking any thing in itself as necessary, then it follows unavoidably . . . that necessity and contingency do not pertain to . . . the things themselves, because otherwise a contradiction would occur." And from this Kant concludes that they can be "only subjective principles of reason," one of which requires us to seek something necessary for everything given as conditioned, while the other prohibits us from ever assuming "anything empirical as unconditioned" (A616 / B644). In other words, here as elsewhere in the Dialectic, the cure consists in recognizing that seemingly

constitutive principles are actually merely of regulative use. As Kant puts it at one point:

The ideal of the highest being is, according to these considerations, nothing other than a **regulative principle** of reason, to regard all combination in the world **as if** it arose from an all-sufficient necessary cause, so as to ground on that cause the rule of a unity that is systematic and necessary; but it is not an assertion of an existence that is necessary in itself. [A619 / B647]

The problem, however, is that for the transcendental realist it is impossible to avoid transposing this regulative principle into a constitutive one, or, as Kant also puts it, “by means of a transcendental subreption, to represent this formal principle to oneself as constitutive, and to think of this unity hypostatistically” (A619 / B647). Actually, what Kant says here is that this subreption is “unavoidable,” which suggests that he is identifying it with the illusion. But since Kant consistently views subreptions as avoidable metaphysical errors, it seems more plausible to take him to be claiming that this subreption, and, therefore, the regulative-constitutive conflation, is unavoidable for the transcendental realist, who is in the grip of the illusion.

At first glance, such a claim might seem bizarre, if not manifestly false. After all, do not many contemporary philosophers, who would gladly characterize themselves as transcendental realists, appeal to something very much like Kant’s regulative-constitutive distinction? To cite a single example with many analogues, consider Daniel Dennett’s well-known conception of an “intentional stance,” which he contrasts with the “physical,” “design” and “personal stances.”<sup>42</sup> A stance for Dennett is simply a predictive explanatory strategy, and the intentional stance is one that treats the behavior of the “system” being investigated as “rational” in the sense that it operates on the basis of beliefs and desires. Since there is no assumption that the system (which might even be a thermostat) really is rational, the idea that it is such is merely a useful fiction, a matter of treating it “as if” it were, which seems very much like a Kantian regulative idea.

This is to ignore, however, a decisive difference between Kantian regulative ideas and the pragmatically justified principles of many contemporary philosophers, namely, that whereas the former are deemed necessary, the latter are assumed to be arbitrary, chosen merely because of their convenience. Moreover, Kantian ideas are necessary in a twofold sense. First, they are necessary products of reason, conceptions to which reason is unavoidably led in its similarly unavoidable endeavor to think the unconditioned. Second, as we shall see in

the following chapter, they are necessary in the further sense that they have an essential cognitive function; that in spite of being “merely regulative,” they are necessary conditions of a coherent use of the understanding. In short, what Kant, in the Dissertation, termed “principles of convenience” (*principia convenientiae*) (ID 2: 418; 414) are not *merely* convenient.<sup>43</sup>

Now while it seems relatively easy for a transcendental realist to posit fictions of the sort used by Dennett, it is by no means clear that such a realist can recognize any such principles in the Kantian sense. Once again, this is because the Kantian conception rests on a sharp distinction between subjective and objective necessity, which is itself a consequence of the transcendental distinction between the sensible and intellectual conditions of cognition. Consequently, lacking the latter distinction, the transcendental realist has no basis for allowing the former.<sup>44</sup> Of course, it remains open for the transcendental realist to deny that there are any subjectively necessary ideas in the first place, which amounts to a complete rejection of the Kantian theory of reason. In that event, however, the realist must either claim that these ideas are objectively necessary, thereby joining forces with the dogmatic metaphysical tradition that Kant is attacking, or provide an alternative account of the rational origins of the central ideas of this tradition. The greatest, perhaps the only, serious attempt to carry out the latter project is Hegel’s *Logic*; but it is doubtful that very many of Kant’s contemporary critics would be tempted to go that route.

## Chapter 15 The Regulative

### Function of Reason

The location of his theory of reason in the Transcendental Dialectic made it difficult for Kant to integrate the positive side of this theory into the framework of the text. Instead, reason appears initially as the “seat of transcendental illusion,” and after a preliminary account of the nature and origin of the transcendental ideas, the great bulk of the Dialectic is devoted to the critical task of undermining the pretensions of transcendent metaphysics. As a result, even though references to the legitimate regulative function of these ideas are scattered throughout the Dialectic and the topic is discussed at some length in the Antinomy, it is only in the Appendix that Kant attempts to deal in a systematic way with the positive function of theoretical reason with respect to the empirical use of the understanding.

Unfortunately, by proceeding in this manner Kant inadvertently created the impression that this account of reason is a mere addendum or afterthought that can be safely ignored by anyone attempting to grasp the basic tenets of the *Critique*.<sup>1</sup> Moreover, if its location in the text were not sufficient incentive to marginalize the argument of the Appendix, additional reasons for doing so are provided by its apparent

contradictions, both within itself and with what Kant supposedly teaches elsewhere in the *Critique*.<sup>2</sup>

These apparent contradictions concern Kant's account of the epistemological status and significance of reason's principle of systematic unity or "systematicity." Kant insists in several places that this principle and its various articulations are merely "logical," that is, expressive of an "interest of reason," without any claim to purchase on the world, while also insisting in other places that they involve a "transcendental presupposition" that must be granted an "objective but indeterminate validity" (A663 / B691). Similarly, at times Kant seems to suggest that these principles have merely a heuristic value, helping to unify our knowledge in a way that the understanding itself cannot achieve, but not contributing materially to its acquisition, while at others he insists on their indispensability for the operation of the understanding. In the latter vein he claims, "[T]he law of reason to seek unity is necessary, since without it we would have no reason, and without that no coherent use of the understanding, and, lacking that, no sufficient mark of empirical truth . . ." (A651 / B679).<sup>3</sup> But since such claims appear to inflate these principles into conditions of the possibility of experience, they are sometimes thought to stand in blatant contradiction to the results of the Analytic. The *Critique*, it is argued, simply has no place for principles that have both a merely regulative status and a genuinely transcendental function.<sup>4</sup>

In recent years, however, the Appendix, or at least its first part, has been viewed in a considerably more favorable light by interpreters interested primarily in Kant's philosophy of science. Although they differ widely in approach and emphasis, these interpreters see the account of systematicity in the Appendix and related discussions in the Introductions to the third *Critique* as an integral component of a Kantian account of the nature and conditions of scientific knowledge.<sup>5</sup> But illuminating as these discussions often are with respect to certain aspects of Kant's theory of empirical science, they tend to neglect the connection of the Appendix with his theory of reason. Consequently, they either ignore entirely or radically downplay both the relation between the regulative function of reason and the doctrine of transcendental illusion and the deeply obscure transcendental deduction of the ideas that Kant provides in the second part of the Appendix.

Given the interests of most contemporary philosophers, this selective approach is understandable; but it also stands in the way of anything approaching an adequate interpretation of the text. Accordingly, the goal of this chapter is to offer a more comprehensive and balanced reading of the Appendix. More



specifically, it will be argued not only that reason has an indispensable regulative function with respect to the acquisition of empirical knowledge, but also that it can have this precisely because (rather than in spite) of the illusory nature of its ideas. In short, transcendental illusion is not only the root cause of metaphysical error, it is also a necessary condition of the successful operation of the understanding.<sup>6</sup> The analysis is divided into three parts: the first two discuss, respectively, the two parts of the Appendix, and the last, which also serves as a conclusion to this book as a whole, considers the connection between Kant's account of the regulative function of reason and transcendental idealism.

## I. TRANSCENDENTAL ILLUSION, SYSTEMATICITY, AND INDUCTION

The complex argument of the first part of the Appendix turns on the relationship between reason's principle of the systematic unity of nature (systematicity) and the idea of a *focus imaginarius*. Since the former is central to Kant's account of the regulative use of reason and the latter provides the metaphor in terms of which he articulates his conception of the transcendental illusion that is inseparable from this use, this is hardly surprising. What is somewhat surprising, however, is the connection between both and the problem of induction. This connection will, therefore, be the central concern of this part of the present chapter. The discussion is divided into four parts. The first analyzes the role of a *focus imaginarius* in what Kant here terms the "hypothetical use of reason," which is equivalent to its inductive use. The second considers the transcendental status of the principle of systematic unity, whose application to nature supposedly requires an appeal to a *focus imaginarius*. The third investigates the epistemic function of this principle and of the sub-principles through which it is expressed (mainly homogeneity and specification). Finally, the fourth discusses its claim, in spite of its illusory nature, to a certain kind of objective validity.

### A. The *Focus Imaginarius* and the Problem of Induction

Kant apparently derived the idea of a *focus imaginarius* from Newton's *Opticks*. Newton's concern was with mirror vision and the optical illusion it involves, whereby an object that really lies behind one's back, and thus outside one's visual field, appears to be in front, just as it would be if the lines of light reflected in the mirror actually proceeded in a straight course.<sup>7</sup> Kant had already made

use this idea in an early work devoted to an ironical critique of metaphysicians in general and Swedenborg in particular (Tr 2: 345–47; 332–34). In the *Critique*, however, he uses it as a metaphor for the transcendental illusion that is supposedly inseparable from the theoretical use of reason. Its purpose is to illuminate the seemingly paradoxical thesis that the “indispensably necessary” regulative use of reason is a function of the illusory nature of its products. This use is initially described as directing the understanding to “a certain goal respecting which the lines of direction of all its rules converge at one point” (A644 / B672). In other words, it is to bring the diverse products of the understanding (its various concepts and empirical generalizations), which in themselves have only a “distributive unity,” to a unity of reason or “collective unity,” in virtue of which they constitute a systematic whole.

Such a project requires a unifying point, which is termed a *focus imaginarius* because it is itself a fiction, a mere idea. As Kant notes, the concepts of the understanding do not really proceed from such a point, “since it lies entirely outside the bounds of possible experience.” Nevertheless, this idea enables the understanding to obtain both the greatest possible unity among its concepts and their extension to the greatest possible range of phenomena. It is by way of illustrating how this works that Kant draws the analogy with the beneficial illusion involved in mirror vision. As he puts it in the crucial last sentence of the paragraph,

Now of course it is from this that there arises the deception, as if these lines of direction were shot out from an object lying outside the field of possible empirical cognition (just as objects are seen behind the surface of a mirror); yet this illusion (which can be prevented from deceiving) is nevertheless indispensably necessary if besides the objects before our eyes we want to see those that lie far in the background, i.e., when, in our case the understanding wants to go beyond every given experience (beyond this part of the whole of possible experience) and hence wants to take the measure of its greatest possible and uttermost extension. [A644–45/B672–73]

Some aspects of the complex metaphor at work in this passage are fairly easy to interpret. Clearly, Kant is suggesting an analogy between our visible and conceptual fields. Just as the mirror image enables the eye to extend the former beyond what is immediately before it, so the idea or intellectual *focus imaginarius* enables the understanding to extend its fund of knowledge beyond its present cognitions, which reflect its limited experience. And just as the former proceeds by producing an optical illusion, so the latter proceeds by producing a transcendental illusion.

Since Kant suggests that what stands in need of explanation is how the understanding is able to go beyond its limited fund of experience and make valid generalizations concerning the whole, it also seems reasonably clear that the problem he has in mind at this point is induction. What remains highly unclear, however, is how the concept of a *focus imaginarius* and the associated transcendental illusion are supposed to help resolve that notoriously intractable problem.

In order to deal with this issue, we must begin with a consideration of the problem of induction as Kant conceived it. Although Hume's classical discussion in the *Enquiry* obviously provides part of the background for Kant's understanding of the problem, a source closer to home may be of greater importance for interpreting the relevance of the mirror-image analogy. This source is Georg Friedrich Meier, who discussed induction in two works with which Kant was quite familiar, namely, *Die Vernunftlehre* (Doctrine of reason) and a volume of excerpts from this larger work, *Auszug aus der Vernunftlehre*, which Kant used as the textbook for his regular logic lectures.<sup>8</sup>

In fact, certain features of Kant's account become wholly intelligible only when seen in this context. Meier includes induction and analogy, which he treats together, among the "mutilated inferences of reason" (*verstümmelten Vernunftschlüssen*), because apart from induction by complete enumeration, which is seldom possible, they lack real legitimacy.<sup>9</sup> Thus, the Baumgartian rationalist is in essential agreement with the Humean skeptic regarding the logical status of inductive procedures. Moreover, Meier's analysis of the problem is also quite similar to Hume's, albeit expressed in different terms. It stems from our epistemic situation regarding empirical cognition. Referring successively to each of our five sensory modalities, Meier points out that they provide access only to "present" objects, not to "absent" ones.<sup>10</sup> Much like Hume, Meier makes the point that any reasoning concerning matters of fact that goes "beyond the present testimony of our senses, or the records of our memory" is both logically and epistemologically suspect.<sup>11</sup>

When Kant's use of the *focus imaginarius* metaphor is seen in this context, its significance becomes fully transparent. The contrast between objects "before our eyes" and those "that lie far in the background" corresponds to Meier's contrast between what is present to the senses and what is absent. Absent objects are likewise phenomena or proper objects of the understanding, but the problem is that, given our epistemic predicament as finite cognizers, we lack access to them. Accordingly, there seems to be no way for the understanding to move from that "part of the whole of possible experience" with which it is contin-

gently acquainted to the far vaster part with which it is not. But, clearly, if the understanding cannot do this, then it cannot make universally valid claims, which, as the “faculty of rules,” is its proper work. Moreover, the problem cannot be resolved simply by appealing to the Analogies. As we have seen, the Second Analogy allows us to infer that every event must have *some* cause, but not to determine what that cause is. Thus, of itself, it is manifestly incapable of grounding induction.

Kant’s own brief discussions of induction, broadly construed to include reasoning by analogy as well as induction proper, which are contained in the various versions of his logic lectures and the *Reflexionen* related to them, agree completely with this analysis.<sup>12</sup> Indeed, apart from the changes brought about by the introduction of the distinction between the determinative and reflective functions of judgment in the third *Critique*, these discussions, stemming from the early 1770s to the late 1790s, reveal a remarkable uniformity in Kant’s conception of the problem. The dominant metaphor is that of induction as “a crutch,” albeit an indispensably necessary one.<sup>13</sup> It is a crutch because the understanding, of itself, is unable to move beyond the given sensory data; it is an indispensably necessary one because, as already noted, being able to do so is a condition of the understanding’s ability to fulfill its proper task. Correlatively, by way of indicating their logically suspect credentials, Kant repeatedly refers to induction and analogy as mere “logical presumptions” rather than genuine inferences of reason, because they lack true necessity.<sup>14</sup>

Interestingly enough, in a late version of his logic lectures, Kant is cited as claiming that “[n]o logician has yet developed analogy and induction properly. This field still lies open” (LD-W 24: 772; 504). Nevertheless, Kant does deal with the topic, albeit as transcendental philosopher rather than as logician, in both the Appendix to the Dialectic and the Introductions to the third *Critique*.<sup>15</sup> In the former, with which alone we are at present concerned, he does so in terms of the distinction between the “apodictic” and the “hypothetical” uses of reason. The former amounts to its use in deductive inference, where the universal is in itself certain and judgment subsumes particulars under it. The latter corresponds to its inductive use, where the particular is certain (the given data) and the universal is assumed only problematically (since it goes beyond these data). In agreement with his other treatments of the topic, Kant remarks:

The hypothetical use of reason, on the basis of ideas as problematic concepts, is not properly **constitutive**, that is, not such that if one judges in all strictness the truth of the universal rule assumed as a hypothesis thereby follows; for how is one to know all

possible consequences, which would prove the universality of the assumed principle if they followed from it? Rather, this use of reason is only regulative, bringing unity into particular cognitions as far as possible and thereby approximating the rule to universality. [A647 / B675]

Assuming that 'approximating' is to be understood as providing an analogue or functional equivalent to the strict universality assumed in deductive reasoning, the obvious question is: What enables empirical generalizations, based on merely a sample of the possible evidence, to play such a role, or, in terms of the operative metaphor, what empowers one to see what lies behind one's back? Clearly, this is an important question, since it goes to the heart of the so-called hypothetico-deductive method of science, which is presumably what Kant had in mind in framing the issue as he did. It is also a *normative* question, since it concerns the grounds by which such a procedure can be justified, rather than, as in the case of Hume's appeal to custom or habit, the means by which it is carried out. Expressed in Kant's own terms, it concerns the justifying grounds for such "presumptions of reason."<sup>16</sup>

Kant's answer to this question turns on reason's principle of systematicity. Although we shall see that Kantian systematicity is much broader in scope than Hume's uniformity principle, since it also encompasses the taxonomic organization of nature into genera and species (an issue that Hume ignores), for present purposes at least they may be taken as functional equivalents. As such, the function of systematicity is to license the inference from something observed to something unobserved, which is also the concern of the hypothetical use of reason. Kant's claim is that "[t]he hypothetical use of reason is . . . directed at the systematic unity of the understanding's cognitions, which, however, is the touchstone of truth for its rules" (A647 / B675).

By characterizing systematic unity as "the touchstone of truth" for the rules of the understanding, Kant appears to be at least gesturing toward some version of coherentism, that is, toward the idea that the criterion of the truth of an empirical proposition is its coherence with the formal and material conditions of experience.<sup>17</sup> Nevertheless, in the present context at least, such a reading is somewhat misleading, since the significance of the claim that systematic unity serves as the touchstone of truth lies elsewhere.<sup>18</sup>

In keeping with the governing analogy, we must take the rules of the understanding to which Kant refers to be experientially derived lawlike statements. Consequently, the principle of the systematic unity of nature provides grounds for assuming that the ordering of phenomena on which we base such statements holds of nature as a whole or, equivalently, of all possible experience.

Once again, however, the problem is that this whole can never be given to us as such, since we can never be in a position to attain the synoptic God's-eye view of things, which would be required to grasp it. That is why Kant insists that systematic unity is merely a "projected unity" (A647 / B675), that is, one which must be assumed as a condition of the operation of the understanding, but which, unlike the unity of phenomena in space and time (which is determinable *a priori*), can never be assured. It is also why a *focus imaginarius*, which functions as ideal ground of this projected unity, is indispensable to reason in its hypothetical use. To view the phenomena of nature as systematically unified in a way that is capable of supporting inferences from part to whole or the observed to the unobserved, just is to view them as deriving their order from such an ideal ground, which we might think of as a superhuman intelligence.<sup>19</sup> This ground is, of course, a fiction, which is why Kant terms it a *focus imaginarius*, but its supposition is nonetheless essential to the extension of the understanding beyond what is immediately given in experience.

Considered in light of Kant's theory of reason, this result may be viewed as a consequence of the inseparability of  $P_1$  and  $P_2$ . When this issue was first considered in Chapter II in connection with Walsh's "scientifically-minded positivist," who is presumably liberated from any contamination by transcendental illusion (this supposedly being a disease peculiar to metaphysicians), it was suggested that such liberation is itself an illusion. In searching for the conditions of natural occurrences, even the dedicated anti-metaphysician necessarily presupposes that the totality of conditions is in some sense given, because the very concept of a condition leads inescapably to the thought of such a totality. In the present case, the totality of conditions is expressed by the idea of the systematic unity of appearances. Consequently, the presupposition of such unity is a necessary condition of the endeavor to unify phenomena by bringing them under empirical laws. Moreover, this systematic unity must be thought as pertaining to nature as a whole, since there is no such thing as a little bit of it.

This does not mean, however, that it is necessary to assume that, if one looks deeply enough, at the end of the day—completed science—such unity will be completely exhibited. It is rather that no amount of recalcitrant experience can be taken to disprove or falsify its presence, since that would mean the end of the understanding.<sup>20</sup> In this respect it differs in kind from any empirical hypothesis, which is certainly at least part of what Kant had in mind in giving the principle of systematic unity transcendental status. As he cryptically puts it, "[H]ere reason does not beg but commands, though without being able to determine the bounds of this unity" (A653 / B681).

## B. Systematicity as Logical and Transcendental Principle

Assuming the necessity of this illusory principle, there remain at least two issues regarding the overall consistency of Kant's account: (1) Is the Appendix itself consistent on the question of the status of this principle, or, as many have claimed, is Kant even there of two minds about the matter? (2) Is the attribution of transcendental status to a principle of reason compatible with Kant's sharp constitutive-regulative distinction?

With regard to the first issue, the problem is clearly indicated in a passage in which Kant comments on the previously noted connection between the hypothetical use of reason and the systematic unity of the understanding's cognitions. Kant begins his discussion by remarking, "From this . . . one sees only that systematic unity . . . is a **logical** principle, in order, where the understanding alone does not attain to rules, to help it through ideas, simultaneously creating unanimity among its various rules under one principle . . . as far as this can be done" (A648 / B676). But, Kant continues:

[W]hether the constitution of objects or the nature of the understanding that cognizes them as such are in themselves determined to systematic unity, and whether one could in a certain measure postulate this *a priori* without taking into account such an interest of reason, and therefore say that all possible cognitions of the understanding (including empirical ones) have the unity of reason, and stand under common principles from which they could be derived despite their variety: that would be a **transcendental** principle of reason, which would make systematic unity not merely something subjectively and logically necessary, as method, but objectively necessary. [A648 / B676]

Although the clear impression left by this passage is that the principle of systematic unity is merely logical, Kant does not deny that it is transcendental. He says merely that *up to that point* (the analysis of the hypothetical use of reason) no reason has been given for assuming that it is anything more than a logical principle expressing an interest of reason. Clearly, this does not preclude claiming later, on the basis of further reflection, that it is really something more, which is just what Kant does.

Even granting this, however, we are still left with our second and more serious question: How, consistently with critical principles, particularly the constitutive-regulative distinction, can Kant claim that the principle of systematic unity is both logical and transcendental? The answer to this question is more complex and consists of two parts.

To begin with, it must be insisted there is no inconsistency in claiming that a principle is both logical (as expressing a necessary interest of reason) and transcendental. The reason the opposite is often thought to be the case stems from the misguided assumption that 'logical' and 'transcendental', like 'constitutive' and 'regulative', refer to mutually exclusive features of principles.<sup>21</sup> Accordingly, on this view, if a principle is merely logical it must also be regulative, which is supposedly incompatible with its also being transcendental.

This widespread misunderstanding results from the combination of a mistaken identification of the regulative with the merely heuristic or optional and an extremely narrow conception of the transcendental as coextensive with the constitutive conditions of possible experience presented in the *Analytic*.<sup>22</sup> Although Kant sometimes labels them "heuristic," we have seen that the transcendental ideas, to which he assigns a regulative function, are themselves claimed to be necessary in the twofold sense of being products of reason following its own principle and of being indispensable for the proper functioning of the understanding. It is the latter point that Kant emphasizes in the Appendix and that accounts for the transcendental status assigned to the principle of systematic unity in spite of its merely regulative function. As Kant succinctly puts it in his discussion of the project of unifying mental phenomena by deriving them from a postulated fundamental power of the mind,

[I]t cannot even be seen how there could be a logical principle of rational unity among rules unless a transcendental principle is presupposed, through which such a systematic unity, as pertaining to the object itself, is assumed *a priori* as necessary. For by what warrant can reason in its logical use claim to treat the manifoldness of the powers which nature gives to our cognition as merely a concealed unity, and to derive them as far as it is able from some fundamental power, when reason is free to admit that it is just as possible that all powers are different in kind, and its derivation of them from a systematic unity is not in conformity with nature? [A650–51 / B678–79]

Here Kant is suggesting that in order to apply the principle of systematic unity, viewed as a logical principle, to nature, it is necessary to assign it a transcendental status and therewith some kind of objectivity. Although this may still seem paradoxical, it is a direct consequence of the doctrine of transcendental illusion, which constitutes the second part of the answer to the second of the two questions posed above.<sup>23</sup> We have already seen that such illusion consists in taking the subjective demand to seek unity (which, as a unity of reason, must be systematic) to reflect an objective necessity pertaining to the things them-



selves (A297 / B353). But since this is precisely what is being done when the logical principle is taken as transcendental, the latter must be seen as the expression of an underlying transcendental illusion.

In order to understand this properly, however, it is essential to keep in mind the distinction between illusion and error. Seen in this light, taking the principle of systematic unity as transcendental is not an error to be avoided by due diligence to its merely regulative status, but a necessary condition of its use by the understanding. Accordingly, what Kant is claiming here is that the slide from the subjective-logical to the objective-transcendental understanding of the principle of systematic unity, like the slide from  $P_1$  to  $P_2$  that it exemplifies, is not only unavoidable but "indispensably necessary." In other words, the very same unavoidable illusion, which if unchecked by transcendental criticism leads reason astray, also turns out to be essential to its proper regulative function with respect to the understanding. As already indicated, this is the deep lesson of the *focus imaginarius* metaphor and of the Appendix as a whole.

### C. The Epistemic Function of Systematicity and Its Forms

Although this may remove the worry about the consistency of Kant's account, both within the Appendix and with the basic tenets of the *Critique*, it still leaves both the transcendental labor supposedly performed by the principle of systematic unity and its claim to a certain sort of objective validity largely unexplained. The former will be discussed in the present section and the latter in the next. In order to deal with these important questions, however, it is necessary to consider Kant's analysis of the logical principles of homogeneity, specification, and affinity or continuity, which are the three forms in which systematicity expresses itself. Unfortunately, the complexity and the contentious nature of the issues once again make it impossible to provide more than a sketch of Kant's position. Nevertheless, this should suffice to bring out the major points.

To begin with, since Kant seems to have viewed affinity or continuity as a combination of the first two (homogeneity and specificity) rather than as an independent principle expressing an additional interest of reason, we shall set it aside here. To simplify things further, the discussion will be limited to the role of these logical principles in the formation of empirical concepts. Although such a focus might seem peculiar, given the preceding emphasis on the problem of induction, which concerns lawlike generalizations rather than the taxonomy of concepts, the two issues are intimately related.<sup>24</sup> As Hannah Ginsborg has

pointed out, without assuming something like natural kinds, one could not even begin to look for empirical laws or hope to distinguish such laws from contingent regularities, which may have a certain generality but not strict universality. Correlatively, determinate empirical concepts presuppose known causal laws, since the properties on the basis of which we classify things prominently include causal properties.<sup>25</sup> Thus, in understanding the role of these principles in the formation of empirical concepts, we shall also be going a long way toward understanding how they enable the human mind to extend its fund of knowledge beyond its present cognitions.

As Kant presents them, the “laws” of homogeneity and specificity represent complementary interests of reason: the former in unity, the latter in differentiation. The basic idea is that the operation of the understanding requires both. Without unity, that is, without the possibility of grouping diverse phenomena into genera and these into higher genera, and so forth, the understanding could gain no foothold in the world. Similarly, without the capacity to draw distinctions within these genera, that is, to divide them into species, and these into subspecies, and so forth, the understanding would be unable to take a single further step. In short, the dual empirical function of the understanding consists in unifying given data by bringing them under empirical concepts and distinguishing between what is unified by introducing additional concepts. But while the process of unification culminates naturally in a highest genus, that of specification proceeds asymptotically toward the goal of complete determination discussed in the preceding chapter. The principle of systematic unity comes into play because this hierarchical organization of empirical concepts is necessarily viewed as a “logical system” allowing inferences from genus to species, and vice versa.<sup>26</sup>

These considerations should suffice to explain why Kant attributes such importance to these principles, viewing them as indispensable for the operation of the understanding, rather than as mere desiderata of reason. Our present interest, however, lies in Kant’s claim that, in spite of their merely regulative function, they have a kind of objective validity.

To begin with, Kant is quite explicit about their transcendental status. Thus, with respect to the first he writes:

The logical principle of genera . . . presupposes a transcendental one if it is to be applied to nature. . . . According to that principle, sameness of kind is necessarily presupposed in the manifold of possible experience . . . because without it no empirical concepts and hence no experience would be possible. [A654 / B682]

Similarly, with regard to the principle of specificity:

But it is easy to see that this logical law would be without sense or application if it were not grounded on a transcendental **law of specification**, which plainly does not demand an actual **infinity** with regard to the variety of things that can become our objects . . . but it does impose on the understanding the demand to seek under every species that comes before us for subspecies, and for every variety smaller varieties. For if there were no lower concepts, then there would also be no higher ones. [A656 / B684]

Having already seen that Kant attributes transcendental status to the principle of systematic unity and that the principles of homogeneity and specificity are expressions of this unity at the empirical level, it is not surprising that Kant likewise gives them a transcendental status. Moreover, once again the key to this status lies in the underlying doctrine of transcendental illusion or, more precisely, the inseparability of  $P_1$  and  $P_2$ . In the present context, this means that we cannot act on these “logical laws” or “maxims” to seek unity and diversity, respectively ( $P_1$ ), without assuming that they are there to be found ( $P_2$ )—not that they ever will be found. In short, this transcendental presupposition is the application condition of these and any regulative principles of reason. Consequently, like the Humean principle of the uniformity of nature, it cannot be viewed as a tentative hypothesis to be tested by the data, because it is a presupposition apart from which there can be no hypotheses to test. As before, the point is simply that “reason does not beg but commands.”

#### D. Systematicity and Objective Validity

It seems to be for this reason that Kant claims that these principles, as synthetic *a priori*, have an “objective but indeterminate validity” (A663 / B691).<sup>27</sup> This is perplexing, because the argument up to this point has affirmed the necessity of proceeding in the investigation of nature *as if* these principles held of the world. But to grant them objective validity, even of an indeterminate or vague sort, seems to suggest that they actually do hold. And if this is correct, then not only the doctrine of transcendental illusion but also the distinctions between understanding and reason, the constitutive and the regulative, on which the whole Dialectic turns, once again appear to be threatened.

As his highly qualified language indicates, Kant was deeply worried by this implication and endeavored to deny it. At the same time, however, his use of the language of objective validity at all suggests that *at this time* he did not recognize any alternative. If, as the argument has shown, these principles are syn-

thetic *a priori*, then their indispensability entails their objective validity. Accordingly, Kant viewed his task as explaining what sort of validity this can be.

Kant's explanation turns on the claim that the principle of systematic unity serves as the "analogue of a schema." The analogy here is purely functional and is based on the premise that the understanding is the proper object of reason. As he puts it, "The understanding constitutes an object for reason, just as sensibility does for the understanding" (A664 / B692). We have seen that the function of sensible schemata is to specify the application conditions of concepts of the understanding. Since reason in its empirical use is directed toward the understanding, the suggestion is that it may also be viewed as supplying such conditions, albeit in a quite different sense. It does so by projecting the idea of "the maximum of the division and unification of the understanding's cognitions in one principle" (A665 / B693). This principle is an application condition of the understanding in the sense that it is indispensable to the latter's endeavor to form empirical concepts and to make lawlike generalizations. Consequently, it may be said to hold of its "object," namely, the understanding. As non-sensible, however, it can be merely the analogue of a schema (in virtue of its functional role), not an actual schema, and, therefore, it does not determine the empirical objects themselves. Nevertheless, since "every principle that establishes for the understanding a thoroughgoing unity of its use *a priori* is also valid, albeit only indirectly, for the object of experience," Kant thinks he is entitled to claim that

the principles of pure reason will also have objective reality in regard to this object, yet not as to **determine** something in it, but only to indicate the procedure in accordance with which the empirical and determinate use of the understanding in experience can be brought into thoroughgoing agreement with itself, by bringing it **as far as possible** into connection with the principle of thoroughgoing unity; and from that it is derived. [A665–66 / B693–94]

Nevertheless, by interjecting the loaded critical notions of objective validity and even objective reality into his account of the regulative function of principles of reason, Kant succeeded in confusing generations of readers and occasioning frequent charges of self-contradiction.<sup>28</sup> As was suggested earlier, the reason for this lies in the fact that Kant did not then (neither in 1781 or 1787) recognize any other way to account for the synthetic *a priori* status of these principles, which is itself a consequence of their indispensability.<sup>29</sup> But, although the issue cannot be pursued here, it should at least be noted that he did find an alternative in the third *Critique* in 1790, with the conception of reflective judgment. For by assigning the legislation of these principles to judg-

ment rather than to reason, Kant was able to preserve their transcendental status without having to attribute to them any sort of objective validity.<sup>30</sup>

## II. THE TRANSCENDENTAL DEDUCTION OF THE IDEAS

The second part of the Appendix is entitled “On the final aim [*Endabsicht*] of the natural dialectic of human reason” (A669 / B697). As the discussion indicates, this title reflects Kant’s underlying teleological assumption that, in spite of the errors and confusion brought about by their misuse, the ideas of reason must have a positive use. What is particularly interesting about this title, however, is that it refers not simply to the final end of human reason but to its “natural dialectic” as well. Thus, the clear implication is that both the transcendental illusion by which reason is haunted and the dialectical perplexities into which it falls on the basis of this illusion themselves serve positive functions. The former is essential to the legitimate regulative use of reason, while the latter, by awakening reason from its “dogmatic slumber,” leads to the recognition of the necessity of a critique of pure reason.

The first of these indispensable services is the topic of this part of the Appendix, where Kant turns from a consideration of the principle of systematic unity and the various forms in which it is expressed, in relation to the empirical use of the understanding, to the official transcendental ideas (soul, world, and God), which were the focus of concern in the main body of the Dialectic. The heart of this discussion is the so-called transcendental deduction of these ideas, which Kant suggests constitutes “the completion of the critical business of pure reason” (A670 / B698).

What supposedly makes such a deduction necessary is the *a priori* status of these ideas, which Kant once again takes as entailing a certain kind of objectivity. Although he readily admits that such a deduction cannot be of the same kind as that of the categories, Kant also insists that “if they are to have the least objective validity, even if it is only an indeterminate one, and are not to represent merely empty thought-entities (*entia rationis ratiocinantis*), then a deduction of them must definitely be possible” (A669–70 / B697–98). Accordingly, the aim of the deduction is to show that the ideas are not merely arbitrarily concocted concepts but have their basis in the very nature of human reason. Or, in Kant’s scholastic terminology, it is to show that they are *entia rationis ratiocinatae* rather than *entia rationis ratiocinantis*, that is, beings of reason functioning rationally rather than of reason functioning sophistically.

Following his usual procedure, Kant first sketches the proposed method of the deduction and then carries it out. As he makes clear at the beginning, the deduction turns largely on the distinction between something being given to reason as an object absolutely and being given only as an object in the idea. The former would require, *per impossibile*, that the object itself actually be given; whereas the latter requires merely that the idea serve as a schema, through which other objects are represented indirectly in their systematic unity by means of their relation to this idea (A670 / B698). As we saw in the preceding chapter, to posit an object in the idea is equivalent to “realizing” the idea, that is, providing it with an intentional object, as contrasted with hypostatizing it, which consists in attributing a real, extra-mental rather than a merely intentional existence to its object.<sup>31</sup> Although Kant does not return to the phrase, he is now suggesting that it is precisely by providing the requisite *focus imaginarius* that an object given in the idea enables the understanding to represent *other* objects indirectly, namely, those lying far in the background, which are otherwise inaccessible. Accordingly, the transcendental deduction of the ideas consists in showing that they have objects in this sense, which presumably would give them a kind of “objective validity.”

It is clear from this that the deduction is crucially dependent upon the account of systematic unity provided in the first part of the Appendix. The basic point is that the transcendental ideas must be seen as indispensable to the exercise of the itself necessary regulative function of reason, which gives them an indirect relation to the understanding and an even more indirect one to the proper objects of the latter, that is, appearances. Some confusion is introduced, however, by Kant’s characterization of these ideas as schemata. Having already claimed that the principles of homogeneity and specificity function as analogues of schemata, by referring to the transcendental ideas as themselves schemata for representing other objects “in accordance with their systematic unity,” Kant now seems to be suggesting that an idea of reason is something like the “schema of a schema,” which seems either to be sheer nonsense or to threaten an infinite regress.

Although this is, in effect, what Kant is claiming, it does not have the deleterious consequences suggested above. Let us keep in mind that in characterizing the principles of homogeneity and specificity as analogues of schemata, he was referring to their function as essential guides to the understanding in its attempt to extend the scope of its empirical knowledge beyond what is immediately present to the senses. What he is now claiming is that by providing the requisite *focus imaginarius*, the transcendental ideas serve as application condi-

tions for these principles of systematic unity. Consequently, in that sense at least, they may be said to function as schemata of schemata, an expression which, to my knowledge, Kant himself never uses.

Before proceeding to the actual deduction of the three ideas, Kant illustrates the point by means of the idea of a supreme intelligence.<sup>32</sup> As an example of an object given merely in the idea, there is no assumption that the world actually was designed by such an intelligence. It is rather that we are enjoined to investigate nature *as if* it were so designed, because doing so is the formula for obtaining the greatest possible unity in our empirical knowledge. As Kant puts it, we are to consider this idea only as a “heuristic” and not as an “ostensive” concept, since “it shows us not how an object is constituted but how, under the guidance of that concept, we ought to **seek after** the constitution and connection of objects of experience in general” (A671 / B699). The task of the deduction is to show that each of the transcendental ideas plays such a role, which is to say that it is indispensable as a *focus imaginarius*.

Clearly, the prospects for such a deduction depend largely upon two factors: (1) establishing the necessity of a *focus imaginarius*; and (2) showing that the transcendental ideas (and only these ideas) are capable of functioning as such. Both of these, however, are far from obvious.

With regard to the first point, even granting everything that Kant has claimed hitherto, this necessity might be questioned. Thus, one might admit (at least for the sake of argument) that the idea of systematic unity is itself necessary for the reasons given in the first part of the Appendix; one might even admit that something like a Kantian *focus imaginarius*, much like a Dennettian “stance,” is often *useful* and can serve an important heuristic function. Nevertheless, it will be objected, this is still far less than is required for a *transcendental* deduction, even of the weak sort appropriate for ideas.

The problem is confounded by the fact that some of what Kant says, particularly in the last pages of the Appendix, seems to support such a line of criticism. For example, after claiming that the idea of a purposive unity, which involves viewing nature as a system of ends, is “inseparably bound up with the essence of our reason” and, therefore, legislative for it, he suggests that it is “very natural to assume a corresponding legislative reason (*intellectus archetypus*) from which all systematic unity of nature, as the object of our reason, is to be derived” (A694–95 / B722–23). And somewhat later he suggests that “we can allow certain anthropomorphisms, which are expedient [*beförderlich*] for the regulative principle we are thinking of, without fear or blame” (A697 / B725). Finally, in affirming the permissibility of considering the purposive orderings

of nature as derived from a divine will, Kant remarks that this is acceptable only if it does not matter whether one says that “ ‘God has wisely willed it so’ or ‘Nature has wisely so ordered it’ ” (A699 / B727).

By claiming that the idea of an archetypal intellect is “natural,” that “certain anthropomorphisms” are “expedient,” and that claims about a divine will are logically equivalent to claims about the wisdom of nature, Kant seems to be suggesting that the idea of a supreme being is less than necessary for the regulative use of reason. Moreover, if this holds in the case of the theological idea, where the claim for a certain kind of necessity appears to be strongest, the problem would seem to be even more acute with regard to the other transcendental ideas, to which Kant devotes considerably less attention.

Although such passages as these point to obscurities in Kant’s account and perhaps to a certain indecisiveness regarding the kind of necessity attributable to the transcendental ideas, they do not threaten his basic claim. To begin with, it is important to keep in mind the context in which these passages occur. Specifically, they are not part of the deduction, which is concerned with the connection between these ideas and the principle of systematic unity. Rather, they are located in Kant’s discussion of purposive unity, which he characterizes as the “highest formal unity” (A686 / B714) sought by reason, but which is not equivalent to the broader notion of systematic unity.

At least in the first *Critique*, the idea of purposive unity, which involves conceiving of nature as embodying a system of purposes (a teleological order) rather than merely a system of laws and an organization in terms of natural kinds (nomological and taxonomic orders) is, from the point of view of the empirical use of the understanding, more like “icing on the cake” than a necessary condition of its operation.<sup>33</sup> In other words, unlike the governing idea of systematic unity, the idea of a purposive unity supervening upon it is not claimed to be the *sine qua non* of *any* extension of the understanding beyond immediate experience. In spite of regarding this idea as an expression of the highest interest of theoretical reason, Kant’s epistemic claims for it are somewhat more modest, namely, that it permits a greater extension of the understanding than would otherwise be possible and that, if properly applied, it can never lead the understanding astray.<sup>34</sup> Thus, by introducing the new concept of a purposive unity, Kant has changed the subject in a significant way, which accounts for the differences in the claims that are made.

This still leaves us, however, with the problem of the necessity of the specific transcendental ideas with respect to the presupposition of a systematic unity. If there is no real difference between thinking that God has wisely willed x and



that nature “wisely” ordered it, it might be thought that we can dispense with the idea of God altogether and simply appeal to that of a systematic unity. From the strictly methodological point of view, which is presumably that of the deduction, it is not clear that the transcendental ideas are not merely excess baggage. Or, somewhat more charitably, what is to prevent us from viewing them as steps on a kind of Wittgensteinian ladder, which once played important roles but have now become superfluous?

Kant’s answer, insofar as he has one, turns once again on the inseparability of  $P_1$  and  $P_2$ . As Kant puts it at one point, “[R]eason cannot think this systematic unity in any other way than by giving its idea an object, which, however, cannot be given through any experience” (A681 / B709). As we have seen, the function of this purely intelligible object is to provide a *focus imaginarius*. Such an imaginary object is necessary in order to apply the principles of systematic unity because we cannot apply them without *at the same time* assuming the point of view in terms of which such unity is alone conceivable by us. To think of nature as embodying a systematic unity and to view it *as if* it were ordered by a supreme intelligence are not two distinct mental acts (in that case they would be separable); rather, to think of it in the former way just is to view it in the latter. In other words, the transcendental ideas are the forms of the thought of the systematic unity of experience, just as the categories are the forms of the thought of its synthetic unity. Moreover, in both cases their transcendental deduction consists in demonstrating this. The difference is that the thought embodied in the ideas is merely regulative, whereas that involved in the categories is constitutive of experience.

Nevertheless, even admitting the viability of this analysis for the theological idea, its applicability to the other two transcendental ideas might still appear dubious. Thus, we shall conclude the present discussion by considering briefly the ideas of the soul and the world, which Kant himself treats in an extremely perfunctory manner in the Appendix, devoting only a single paragraph to each.

Of these, the idea of the soul is at once the most interesting and, from a contemporary point of view, most problematic. Indeed, few contemporary philosophers, including Kantians, would countenance granting regulative status to the idea of an immaterial substance, which is just what Kant does.<sup>35</sup> As before, the concern is with the conditions of the thought of a systematic unity, this time of the appearances of inner sense. The claim is simply that the idea of the soul is necessary for the thought of such a unity, because the latter can be thought only

by considering all determinations as in one subject, all powers, as far as possible, as derived from one fundamental unique power, all change as belonging to the states of one and the same persisting thing, and by representing all **appearances** in space in space as entirely distinct from the actions of **thinking**. [A682–83 / B710–11]

For Kant at least, the last of these ways of conceiving the relation of mental states and operations made possible by the idea of the soul is clearly the most significant. In other words, Kant's main reason for conceiving of the mind in the light of the idea of the soul appears to be that it blocks the intrusion of physicalistic explanations into the psychological domain, and this might, in turn, be claimed to be a necessary condition for a systematic description of psychological phenomena.<sup>36</sup> As Kant puts it,

For then empirical laws of corporeal appearances, which are of an entirely different species, will not be mixed up in the explanation of what belongs merely to **inner sense**; then no windy hypotheses about the generation, destruction or palingesis of souls, etc., will be admitted; a consideration of this object of inner sense as a whole will therefore be instituted; and this will not be mixed up with properties of any different kind. . . . [A683 / B711]

Although the “windy hypotheses,” which Kant wished to excise from the study of the mind, are clearly quite different from those that would be of interest to contemporary philosophers of mind or cognitive scientists, they share a common feature with many of the latter, which, for want of a better term, we may call a denial of the integrity of the mental. As the quoted passage indicates, the prime theoretical function of the idea of the soul is to preserve this integrity at the explanatory level. Naturally, this pales in comparison with its practical function, with which Kant was far more deeply concerned, but it is not for that reason trivial.

Moreover, in virtue of its rejection of the intrusion of any reductive physicalistic explanations into the analysis of thinking, Kant's account bears a certain similarity to the prevalent contemporary approach termed “functionalism.”<sup>37</sup> Nevertheless, there remains a crucial, though not always recognized, difference between the idea of the soul operative in the Kantian scheme and the computer-inspired hardware-software metaphor underlying present-day functionalist accounts. For whereas the latter is wedded to a naturalistic, broadly materialistic (though non-reductive) view of mental processes, according to which they are regarded as causally conditioned events in virtue of an assumed token-token identity with their physical counterparts, the Kantian idea of the soul allows space for a genuine spontaneity, which is essential to Kant's conception of

cognition and, therefore, the argument of the *Critique* as a whole.<sup>38</sup> And given such a program in psychology, it does not seem far-fetched to claim that something like the idea of the soul is required as a regulative principle. At the very least, it seems as well suited to serve as a *focus imaginarius* as the computer-based model to which contemporary functionalists appeal.

Due at least in part to the structure of the Dialectic, Kant's discussion in the Appendix of the regulative function of the cosmological idea (or ideas)<sup>39</sup> is, if anything, even more perfunctory and less satisfying than his treatment of the psychological idea. As Kant has already discussed the regulative functions of these ideas at some length in the last section of the Antinomy chapter, there appears to be very little left to be said on the matter. Thus, Kant quite reasonably refers back to that account for a fuller discussion of this complex issue, and he is here content to remind the reader that, on the one hand, "in the explanation of given appearances . . . we ought to proceed as if the series were in itself infinite, i.e., proceed *in indefinitum*," and, on the other hand, with respect to freedom, "we should proceed as if we did not have before us an object of sense but one of pure understanding, where the conditions can no longer be posited in the series of appearances, but are posited outside it, and the series of states can be regarded as if it began absolutely through an intelligible cause . . ." (A685 / B713). But since, in the latter case the function of the idea is practical rather than theoretical, regulating our conception of ourselves as rational agents, it falls outside the scope of our present concerns.<sup>40</sup>

Nevertheless, the structure of the Dialectic does not of itself explain the unsatisfactory nature of the Appendix's account of the regulative function of the cosmological idea. As Kant himself suggests, the real problem lies in the peculiar nature of this idea. Unlike its psychological and theological counterparts, the idea of the world is allegedly self-contradictory. Indeed, we have seen that this is precisely why it generates antinomies. Moreover, this suggests that we cannot so readily assume it to be "objective and hypostatic" (A673 / B701). But if we cannot do this, it becomes difficult to see how it could have an object, even "in the idea," which, in turn, appears to be required if it is to function as a *focus imaginarius*.

Although Kant does not deal explicitly with the issue, it seems possible to piece together a response from a consideration of what he says in the Antinomy chapter regarding the regulative functions of the distinct expressions of cosmological idea operative in the four Antinomies. At least with respect to the Mathematical Antinomies, to which the present discussion must be limited, Kant's position seems to be that the regulative work is done entirely by the ideas in the

forms of which they are understood in the antitheses, that is, as ideas of infinite totalities of one kind or another. Since these forms (or ideas) are illusory, but not self-contradictory, they can function regulatively.

In each of the Mathematical Antinomies, the idea may be seen to function in this way by providing a *focus imaginarius*, which serves to prod the understanding to search continually for ever new conditions, on the grounds that at no given point in the regress may it be claimed that the totality thereof has been attained. Accordingly, here as elsewhere in the Dialectic,  $P_1$  and  $P_2$  are at work: the former requiring the understanding to search for conditions and the latter preventing it from stopping at any point prior to the (impossible) attainment of their totality.

Nevertheless, there is a significant difference between the way in which the idea of an infinite totality functions regulatively as a *focus imaginarius* in each of the Mathematical Antinomies. Although Kant glosses over this difference in the Appendix, merely suggesting, as we have already seen, that “in the explanation of given appearances . . . we ought to proceed as if the series were in itself infinite, i.e., proceed *in indefinitum*,” in the Antinomy chapter itself he distinguishes sharply between a regress that proceeds merely *ad indefinitum* and one that proceeds *ad infinitum*. The former is connected with the resolution of the First Antinomy and the latter with that of the Second.

As the terms suggest, the difference is between a regress to which no limit can be specified and one that must be assumed to proceed to infinity.<sup>41</sup> According to Kant, the former applies to the First Antinomy because the magnitude (age and size) of the world can never be given as a whole (either as finite or infinite). Thus, all that may be said about the regress here is that for every given member of the series of conditions we are required (by  $P_1$ ) to seek out higher and more remote conditions (A519 / B547). By contrast, since in the case of the Second Antinomy (concerning divisibility) the whole is assumed to be given prior to the regress to its parts (otherwise there could be no regress), we must assume that these parts are likewise already given (*in potentia*). Consequently, the principle that the parts are only given in and through the regress does not apply here. Rather, what cannot be claimed to be given prior to the regress is the totality of the *division* of the parts, that is, the decomposition of the whole. This precludes the dogmatic claim of the Antithesis that the whole is composed of an actual infinite number of parts, but it leaves in place the regulative principle that the regress be viewed as proceeding *ad infinitum*.<sup>42</sup>

### III. SYSTEMATIC UNITY AND TRANSCENDENTAL IDEALISM

The connection between Kant's account of the regulative function of reason sketched in the Appendix and his transcendental idealism is our final concern. Although this is potentially a large topic, since most of the ground has already been covered it is possible to be relatively brief. To begin with, Kant's underlying thesis that "reason does not beg but commands" obviously brings with it an idealistic commitment of a sort. The systematic unity of nature is a projected unity, not one that is inferred from a pre-given order of things. Moreover, we have seen that this is essential to its regulative function, since it ensures that the assumption that nature embodies such a unity cannot be disconfirmed by recalcitrant experience. Nevertheless, this does not of itself appear to entail transcendental idealism, understood as the doctrine that we can know things only as they appear (in accordance with the subjective conditions of human sensibility), not as they may be in themselves. Accordingly, in spite of everything that has been said up to this point, it might seem that one could accept Kant's claims about the regulative function of reason, while comfortably remaining a realist in the transcendental sense.

Although this is mistaken, the mistake is not apparent unless one construes transcendental idealism in the methodological or metaphilosophical sense as the antithesis of transcendental realism. Once again, the two forms of transcendentalism must be understood as constituting two all-inclusive and mutually exclusive metaphilosophical alternatives or standpoints.

We have seen repeatedly that transcendental idealism functions in the Dialectic mainly as the antidote to the potentially pernicious influence of transcendental illusion. Although it does not remove this illusion (that being impossible), it does render it harmless by separating it surgically from the transcendental realism with which it is naturally connected and which is the true source of the problem. As was suggested at the end of the preceding chapter, however, this idealism also functions positively by creating the conceptual space necessary for an understanding of the ideas as at once merely regulative and indispensable. Since it cannot recognize any such space, transcendental realism is forced either to deflate this regulative function into a purely heuristic one, which may be adopted or discarded as the occasion warrants, or to inflate the ideas of reason into fully objective principles, having a status equivalent (or even superior) to that of the categories. In fact, such an understanding of the al-

ternative possibilities is continually reflected in the endless debate in the secondary literature about the true status and function of these ideas in the *Critique*.

Clearly, it is this positive function of transcendental idealism that is directly germane to the argument of the Appendix. In particular, even though it is never mentioned, transcendental idealism underlies Kant's use of the mirror analogy to illuminate the problem of induction, since it is such idealism alone that allows for a positive and non-deceptive use of the *focus imaginarius*. Moreover, this claim does not rest on a dubious identification of transcendental realism with naive realism or some close relative thereof. Putting it in terms of Kant's metaphor, the point is rather that transcendental realism, in any of its guises, is simply not equipped to view the objects lying in the background (the absent particulars) *as if* they are present. If they are present at all for such a realist, they must be *really* present, as they presumably are from the God's-eye view that is integral to such realism.<sup>43</sup>

Conversely, if objects are known not to be present, then the transparently fictional assumption of their presence can have nothing more than a provisional, merely instrumental justification, lacking any genuine normative force. By contrast, in regarding the God's-eye view as itself a necessary projection of human reason rather than as a pre-given norm to which human knowledge must conform, transcendental idealism may be said to allow for the possibility of conceiving absent objects as "virtually present," that is, as present in idea, which, in turn, provides a normative ground for the extension of the empirical use of the understanding beyond the given data. In short, the assignment of any normative force to the projections of reason is inseparable from a more radical reconfiguration of normativity, which is just Kant's "transcendental turn," understood as a turn not *to* the transcendental but *from* one form of transcendentalism to another.

Precisely the opposite, however, has been claimed by many anti-idealists with regard to Kant's thought in general and by Ralph Walker with regard to the specific problem of induction. Correctly noting that transcendental idealism is concerned with the formal conditions of experience, Walker claims that for this very reason it founders over the problem of induction. According to Walker, this is because the kind of *a priori* guarantee of the conformity of objects to the sensible and intellectual conditions of cognition supposedly provided by transcendental idealism is not available for the *a posteriori* elements in experience with which induction is concerned.<sup>44</sup> Thus, for Walker, transcen-

dental idealism is unable to handle the problem of induction, by which he apparently understands providing some kind of guarantee that nature will conform to our experientially based hypotheses. In fact, on his view the problem is even worse. Not only is transcendental idealism incapable of dealing successfully with this problem, but since, according to Walker, its very *raison d'être* is to provide such guarantees, it loses its entire motivation. In other words, it is not that the problem of induction shows transcendental idealism to be somehow incoherent; it is rather that, by revealing its inadequacy for that particular task, it removes any incentive to affirm it at all.<sup>45</sup>

This is a strange argument for a number of reasons. The first of these is Walker's apparent assumption that the function of transcendental idealism is to provide some kind of ironclad guarantee that the world conforms *in toto* to our forms of cognition. It is as if transcendental idealism is viable only if it is able to eliminate *all* worries about such conformity, or as if it cannot guarantee conformity anywhere unless it can do so everywhere. But this ignores the merely formal nature of this idealism, which Walker himself emphasizes. Thus, even if Walker were correct in his assessment of the irrelevance of transcendental idealism to the problem of induction, this would not invalidate its claim to account for the possibility of *a priori* knowledge, by providing the "formal conditions of empirical truth." Second, although it is obviously true that accounting for the possibility of synthetic *a priori* knowledge, which does require providing some such guarantee, is a major goal of the *Critique* and depends crucially upon transcendental idealism, this is not the only significant service that the latter performs. As we have seen, Kant assigns equal importance to the resolution of the antinomies, which involves nothing less than avoiding "the euthanasia of pure reason." Finally, Walker apparently assumes that if there were a genuine problem of induction, which he seems to deny, it could be resolved only by providing some such guarantee. This completely disregards, however, the analysis that Kant offers, which turns on the assignment of a normative role to the projections of reason (and later reflective judgment) but offers no such guarantee.

Of course, one may reject the whole Kantian approach to induction, which finds its definitive expression in the third *Critique*. One may also reject transcendental idealism, though, as has been argued in this book, such a rejection comes at a steep price. What one cannot do is separate the two, as if this idealism had nothing to do with Kant's treatment of the problem of induction. Indeed, so long as transcendental idealism is understood to be the only alternative

to transcendental realism, it cannot, here as elsewhere, be neatly separated from the substantive claims of the *Critique*. Thus, in spite of the efforts of many interpreters to have it otherwise, the rejection of transcendental idealism amounts to the rejection of Kant's theoretical philosophy as a whole. That is why the correct interpretation of this idealism is so important.



## Notes

### PREFACE

1. Strawson, *The Bounds of Sense*, esp. pp. 38–42, 253–73.
2. Guyer, *Kant and the Claims of Knowledge*, p. 336; Langton, *Kantian Humility*, pp. 8–12.
3. Grier, *Kant's Doctrine of Transcendental Illusion*.
4. These include those of Brandt, *The Table of Judgments: Critique of Pure Reason A 67–76; B 92–101*; Wolff, *Die Vollständigkeit der kantischen Urteilstafel*; Longuenesse, *Kant and the Capacity to Judge*. Prior to their work, virtually the only commentator to take seriously the supposed completeness of Kant's table of judgment forms was Klaus Reich in *The Completeness of Kant's Table of Judgments*. These recent writers take up Reich's project, while rejecting his analysis.
5. See Allison, *Kant's Theory of Freedom and Idealism and Freedom*.

### CHAPTER 1

1. For a useful discussion of these forms of anti-realism and their connection, or lack thereof, with Kant's idealism, see Karl Ameriks, "Kantian Idealism Today," *History of Philosophy Quarterly* 9 (1992), pp. 329–42.
2. Although I occasionally depart from the procedure, largely, though not entirely, for stylistic reasons, in contrast to the first edition, I here tend to use the

- term ‘cognition’ rather than ‘knowledge’ to render Kant’s ‘*Erkenntnis*’. In addition to the fact that it correlates with the Latin ‘*cognitio*’, which Kant equates with ‘*Erkenntnis*’, this has two advantages, which more than compensate for a certain awkwardness: first, cognition, unlike knowledge, may either succeed or fail to attain its object and, second, it can be used in the plural. In using this term, I am here also following the Guyer and Wood as opposed to the Kemp Smith translation, which I used in the first edition.
3. In the first edition of *Kant’s Transcendental Idealism* I referred to this as the “standard picture.” Fortunately, this description is no longer accurate, though it remains a widely accepted reading.
  4. These include Pistorius, Eberhard, Jacobi, Maimon, and Aenesidemus-Schulze. Perhaps the clearest contemporary expression of this kind of interpretation, however, is to be found in the notorious Garve-Feder review to which Kant himself replied in Pro 4: 372–80; 160–66. For an account of many of these interpretations and criticisms of Kant, see Vaihinger, *Commentar zu Kants Kritik der reinen Vernunft*, vol. 2, pp. 494–505.
  5. Strawson, *The Bounds of Sense*, p. 16.
  6. The above is admittedly an oversimplified account of Strawson’s position, based largely upon his introductory account in *The Bounds of Sense*, pp. 38–42. He also discusses transcendental idealism in several other places in the work, most notably on pp. 235–62, and he distinguishes among various possible interpretations. My present concern, however, is not to examine the details of Strawson’s interpretation and critique but merely to use it as illustrative of what remains a fairly common view of transcendental idealism. I have dealt specifically with Strawson’s views in my “Transcendental Idealism and Descriptive Metaphysics,” *Kant-Studien* 60 (1969), pp. 216–23. For a similar line of criticism, see H. E. Matthews, “Strawson on Transcendental Idealism,” *Philosophical Quarterly* 19 (1969), pp. 204–20.
  7. It should be noted, however, that Hume explicitly affirmed such a doctrine in the *Treatise*, which may give one pause before dismissing it as incoherent; see Lorne Falkenstein, “Hume on Manners of Disposition and the Ideas of Space and Time,” *Archiv für Geschichte der Philosophie* 79 (1997), pp. 179–201. Nevertheless, I shall not discuss this issue here, since I can see little evidence that Kant actually held such a view.
  8. Strawson, *Bounds of Sense*, pp. 235–39.
  9. Prichard, *Kant’s Theory of Knowledge*, pp. 71–100. A rigorous and sensitive critique of Prichard’s Kant interpretation has been provided by Bird, *Kant’s Theory of Knowledge*, esp. pp. 1–17. Although I differ from Bird on a number of issues, the strategy of this chapter, and also the overall direction of my interpretation, owes much to his work. To my mind, he deserves credit for being the first English-language Kant commentator of this generation seriously to challenge the standard picture.
  10. Prichard, *Kant’s Theory of Knowledge*, pp. 78–79.
  11. Strawson, *The Bounds of Sense*, p. 240.
  12. For a critical discussion of some of these attempts, see Walker, *Kant*, esp. pp. 14–23.
  13. Guyer, *Kant and the Claims of Knowledge*, p. 336.
  14. *Ibid.*, p. 333.
  15. *Ibid.*
  16. *Ibid.*, pp. 334–5.

17. Ibid., p. 335.
18. Ibid.
19. Cf. Pro 4: 289; 84; Gr 4: 451–52; 98–99.
20. Jacobi, *Werke*, vol. 2, p. 304.
21. Guyer, *Kant and the Claims of Knowledge*, p. 335.
22. As we shall see in chapter 10, Guyer does at times acknowledge that Kant held to something like a two-aspect view of transcendental idealism, but he sees this as an abandonment in 1787 of the original view of 1781 and claims that this later conception is likewise irrelevant to the central arguments of the *Analytic*.
23. Langton, *Kantian Humility*.
24. For Langton's views on idealism in Kant, see *Kantian Humility*, esp. p. 6 and her last chapter, "Realism or Idealism?" In the latter she acknowledges that Kant does claim that space (and presumably time) are ideal, while minimizing the significance of this concession by denying that the ideality thesis applies to things in space and time, that is, phenomena.
25. Ibid., p. 8.
26. Strawson, *The Bounds of Sense*, p. 250.
27. Langton virtually neglects time, and she has relatively little to say about space. Her justification for this is her conviction that Kant's views about humility are independent of his views about space and time. See especially *Kantian Humility*, p. 102, note 7, and p. 211. I shall challenge this claim in the sequel.
28. See Langton, *Kantian Humility*, pp. 124–25 and 139. For her discussion of supervenience see esp. pp. 79–88.
29. Ibid., p. 50.
30. As we shall subsequently see in more detail, Kant insists on a sharp distinction between a real and a merely logical use of the categories. The former is their use in synthetic judgments, whereas the latter involves merely analytic judgments. In my judgment, one of the underlying weaknesses of Langton's interpretation is her complete neglect of the analytic-synthetic distinction, which, whatever difficulties it may involve, is unquestionably central to Kant.
31. Malte Hossenfelder, "Allison's Defence of Kant's Transcendental Idealism," *Inquiry* 33 (1990), pp. 467–79, esp. pp. 468–69.
32. I am here attempting to correct a defect in my original analysis, where I suggested that transcendental idealism follows from the mere concept of an epistemic condition. In response to this, some critics charged me with ambiguity, noting that I sometimes present transcendental idealism in this way but at other times express the more orthodox view that it depends on Kant's conception of human sensibility as having *a priori* forms or conditions. A very useful contribution to the discussion has been made by Karl Ameriks, who distinguishes between "non-specific" and "specific" versions of transcendental idealism. The former attempts to define transcendental idealism in broad epistemological terms as affirming the dependence of objects on our conceptual schemes, cognitive capacities, theories, or the like. The latter locate the essence of Kantian idealism in his theory of sensibility. (See "Kantian Idealism Today," pp. 333–34.) Put in these terms, the view I am here advocating may be seen as a combination of both. It shares with the for-

- mer versions a focus on the conditions of cognition, which results in an epistemologically rather than a metaphysically oriented idealism; it shares with the latter a focus on the sensible conditions of human knowledge, which it sees as a consequence of its discursive nature.
33. Perhaps the first to challenge this thesis was Salomon Maimon. Maimon's critique of Kant on discursivity is analyzed at length by Peter Thielke, "Discursivity and Its Discontents: *Maimon's Challenge to Kant's Account of Cognition*," Ph.D. dissertation, University of California, San Diego, 1999 (unpublished). More famously, it was rejected by Fichte with his doctrine of intellectual intuition and by Hegel, first in *Glauben und Wissen* with his own version of intellectual intuition, and later in the *Phenomenology* in connection with the conception of absolute knowledge.
  34. Strawson, *The Bounds of Sense*, p. 20.
  35. *Ibid.*, pp. 20–21.
  36. Here I am in general agreement with the analysis of this issue by M. Glouberman, "Conceptuality: An Essay in Retrieval," *Kant-Studien* 70 (1979), pp. 383–408. Nevertheless, I differ from Glouberman on some points and take the analysis in a somewhat different (though complementary) direction. Specifically, I explore the connection between the discursivity thesis and Kant's critique of transcendental realism (chapter 2) and the analytic-synthetic distinction (chapter 4), neither of which he mentions. Perhaps most important, I insist on the importance of this thesis for understanding Kant's idealism (something which he merely hints at near the end of his discussion).
  37. Once again, this is to be seen as a modification of my earlier view that the discursivity of human cognition is treated by Kant as a kind of brute fact for which no argument is advanced. Although I still believe it true that Kant does not explicitly provide such an argument, I think that he gives us the requisite materials. The need for such an argument has been emphasized by Thielke, *Discursivity and Its Discontents*.
  38. The connection between these two defining features of a Kantian sensible intuition will be explored in chapter 4.
  39. Thus, for Kant the concept of an intuitive intellect is intended to model the divine mind. Although he considers the conception of such an intellect problematic, since we have no way to understand its possibility, Kant thinks that it serves an important regulative function, indicating the ineliminable limits of our discursive cognition. Kant's canonical discussion of the relations between these two forms of "intellect" (discursive and intuitive) is in §§ 76 and 77 of the *Critique of Judgment*. The issue will be explored further in chapter 2.
  40. Kant makes this basic point in a number of different ways, suggesting the need for a synopsis attributed to sense as well as for a synthesis of the understanding (A94 / B127 and A97) and that the senses are *determinable* (my emphasis) but not determining (B151–52).
  41. See, for example, O'Shaughnessy, *The Will*.
  42. I analyze Spinoza's account of the mind-body relation in my *Spinoza: An Introduction*, pp. 85–100.
  43. This view has been advanced in a number of papers by Ralf Meerbote, including "Kant

on the Nondeterminate Character of Human Actions,” in *Causality, Freedom, and Objectivity*, ed. by William A. Harper and Ralf Meerbote, pp. 138–63; “Space and Time and Objects in Space and Time: Another Aspect of Kant’s Transcendental Idealism,” in *Minds, Ideas, and Objects*, ed. by Phillip D. Cummins and Guenter Zoeller, pp. 275–90; “Wille and Willkür in Kant’s Theory of Action, in *Interpreting Kant*, ed. by Moltke S. Gram, pp. 69–80. See also Hudson, *Kant’s Compatibilism*.

44. For reasons to become clear in the course of this book, there is an important asymmetry here. The reason for this is that in considering objects as they appear or as appearances, one is actually considering them as subject to intellectual as well as sensible conditions (the schematized categories and the Principles), whereas in considering them as they are in themselves the converse does not hold.
45. Among the many places in which Kant explicitly restricts his denial of the use of the categories (or the pure understanding) with regard to things as they are in themselves or noumena to *synthetic* judgments are A 276 / B 273; A 286 / B 342–43; A 433 / B 461; A 609 / B 663.
46. Langton, *Kantian Humility*, pp. 9–10.

## CHAPTER 2

1. What makes them both forms of transcendentalism is their complete generality, here understood as a concern with the empirical as such.
2. Two interpreters who take this view are Colin Turbayne, “Kant’s Refutation of Dogmatic idealism,” *Philosophical Quarterly* 5 (1955), p. 228, and Sadik J. Al-Azm, *The Origins of Kant’s Argument in the Antinomies*, p. 148.
3. In the *Critique of Pure Reason* (B 53) Kant seems to equate “absolute” with “transcendental” reality. The notion of absolute reality goes back at least as far as the Dissertation, where Kant criticizes the conception of time as something “posited in itself and absolutely” (*in se et absolute positum*) (Diss 2: 401–2; 395). For a discussion of some of these terminological points see Hinske, *Kants Weg zur Transzendentalphilosophie*, esp. p. 49.
4. See Gr 4: 440–45; 89–93 and KrV 5: 39–40; 172–73.
5. I analyze Kant’s systematic opposition between autonomy and heteronomy, conceived as alternative models of volition and the associated claim that other moral theories are committed to the latter model in my *Kant’s Theory of Freedom*, pp. 93–106. My claim there is that heteronomy should be seen as the moral equivalent of transcendental realism, whereas the present claim is just the inverse of this.
6. Cf. Turbayne, “Kant’s Refutation,” and Norman Kemp Smith, *A Commentary to Kant’s “Critique of Pure Reason,”* pp. 301ff.
7. At the previously cited A 369, Kant characterizes transcendental realism as holding that outer appearances (spatial objects) are “outside us [*ausser uns*] according to pure concepts of the understanding. Since to think in accordance with pure (unschematized) concepts is to make what Kant terms a transcendental (i.e., completely universal) use of the understanding, this is equivalent to viewing these appearances as *ausser uns* in the transcendental sense.

8. I am aware that classifying Hume as a phenomenalist is a controversial issue in Hume interpretation. Nevertheless, I believe myself justified in treating him as such for present purposes, since this is clearly how Kant understood Hume.
9. Once again, I would like to remind the reader that it is not a question of the fairness to Berkeley of Kant's reading but of what it reveals about Kant's own views. For my full account of Kant's interpretation of Berkeley, see Allison, "Kant's Critique of Berkeley," *Journal of the History of Philosophy* 11 (1973), pp. 43–63.
10. Kant also makes essentially the same point about Berkeley in connection with the Refutation of Idealism (B 274–75).
11. This analysis is to be contrasted with a discussion of the same text by Glouberman, "Conceptuality: An Essay in Retrieval," pp. 391–92. Rather than focusing on the role of sensibility, Glouberman emphasizes Hume's rejection of conceptual representation as an adequate mode of cognition. Although this is no doubt true, I do not think that it captures Kant's emphasis in the passage in question. Moreover, here it is important to keep in mind that for Kant even the private data of inner sense are given to the mind under the form of time and hence count as appearances. In other words, these data are *in uns* in both the empirical and the transcendental sense.
12. The analysis of this passage is greatly indebted to the discussion of Beck, *A Commentary on Kant's Critique of Practical Reason*, pp. 181–82.
13. This is emphasized by Glouberman, "Conceptuality: An Essay in Retrieval," p. 390 and *passim*. He does not, however, connect this important point with transcendental realism and Kant's critique thereof.
14. See Spinoza, *Ethics*, 2, prop. 40, scholium 1. I discuss this contrast in my *Benedict de Spinoza: An Introduction*, pp. 116–19 and *passim*.
15. A good example of this (noted by Glouberman, "Conceptuality: An Essay in Retrieval," p. 368) is Hume's discussion of "distinctions of reason" (*A Treatise of Human Nature*, pp. 24–25). Hume takes these to be distinctions that the understanding draws between aspects of things that are inseparable in the original impressions, e.g., the distinction between shape and color. Although such distinctions turn out to be crucial for the purposes of communication and science, Hume's critical point is that they are *only* distinctions of reason and, therefore, have no real purchase on the nature of things.
16. Merold Westphal, "In Defense of the Thing in Itself," *Kant-Studien* 59 (1968), pp. 118–41, argues in a similar vein that things as they are in themselves are to be understood as things as they are for God. Although this is certainly correct, I do not believe that Westphal's attempt to anchor transcendental idealism in a theistic metaphysics is particularly illuminating. Kant presumably was a theist (of sorts), but this does not account for his transcendental distinction and limitation of cognition to things as they appear. Moreover, Westphal fails to discuss the issue of the commitment of Kant's predecessors to a theocentric model, and he explicitly downplays the significance of the discursivity thesis. (On the latter point, see *loc. cit.*, p. 131).
17. Far closer to home, Hilary Putnam refers critically to such a model in arguing for his own, Kant-inspired, "internal realism." See *Reason, Truth and History*, pp. 60–64.
18. Cf. Spinoza, *Ethics*, II, prop. 54, corollary 2.
19. The fullest discussion of this point is by Gurwitsch, *Leibniz, Philosophie des Panlogismus*,

- esp. pp. 23–31, 142–51, 450–54. It is also brought out in connection with Kant by Gottfried Martin, who refers to it as the “theological foundation of truth.” *Kant’s Metaphysics and Theory of Science*, p. 62.
20. Leibniz, *New Essays on Human Understanding*, p. 447.
  21. This is shown nicely by Gurwitsch in his discussion of the “Affinität des menschlichen und göttlichen Geistes” (*Leibniz*, pp. 142–44).
  22. Leibniz, *Discourse on Metaphysics*, p. 13.
  23. Cf. A43–44 / B61–62, Pro 4: 290–91; 85–86, UE 8: 219; 310.
  24. For a discussion of this see Allison, *The Kant-Eberhard Controversy*, pp. 75–92.
  25. This is to be contrasted with the account of Kant’s interpretation and critique of Leibniz provided by Langton (*Kantian Humility*, pp. 197–203). According to Langton, Kant’s claim in the *Critique* that Leibniz took appearances for things in themselves is to be understood in metaphysical terms as charging him with simply identifying natural appearances with monads. Kant then allegedly corrects this in his response to Eberhard by stating that Leibniz (in contrast to Eberhard) asserted merely that the monads are the *grounds* of appearances, which is close to the Kantian view. In my view, this completely neglects the essentially epistemic thrust of Kant’s critique of Leibniz and suggests a non-existent conflict between his treatment of Leibniz in the *Critique* and *On a Discovery*.
  26. Locke, *An Essay Concerning Human Understanding*, p. 439.
  27. *Ibid.*, p. 440.
  28. *Ibid.*, p. 45.
  29. *Ibid.*
  30. *Ibid.*, p. 303.
  31. Moltke Gram has suggested that Kant understands the notion of intellectual intuition in three senses: (1) an intellect that can be aware of objects independently of any forms of intuition whatsoever; (2) an intellect that could be aware of the sum total (*Inbegriff*) of appearances; and (3) a creative, archetypal intellect; see “Intellectual Intuition: The Continuity Thesis,” *Journal of the History of Ideas* 42 (1981), pp. 287–334, and *The Transcendental Turn*, p. 223, note 7. Unfortunately, he neglected a fourth sense (or perhaps an aspect of the first), namely, an intellect that can be aware of its object in its full concreteness and particularly, without any dependence on conceptualization. Empiricists like Locke may be said to appeal to intellectual intuition in *at least* the fourth sense (which, given the analysis in the first chapter, is really inseparable from the first).
  32. Consequently, both thinkers deny the discursivity thesis. As Kant puts it, “Instead of seeking two entirely different sources of representation in the understanding and the sensibility, which could judge about things with objective validity **only in conjunction**, each of these great men holds on only to one of them, which in his opinion is immediately related to things in themselves, while the one does nothing but confuse or order the representations of the first” (A271 / B327).
  33. A clear expression of this is to be found in Kant’s positive yet critical remarks about Maimon in his well-known letter to Herz of May 26, 1789 (Herz had sent Kant the manuscript of Maimon’s *Versuch über die Transzendentalphilosophie mit einem Anhang über die symbolische Erkenntnis*). Kant’s basic complaint is that Maimon assumed that human reason is of the same kind as the divine and differs from it merely in degree. Significantly,

- Kant also suggests that the antinomies provide a good antidote to such a view. (See B11: 52–54; 316–18). We shall explore the latter point in chapter 13.
34. Implicit in all of this is the equivalence of ‘form’ and ‘condition’. This point is discussed in chapter 5 and elsewhere. For a detailed account of Kant’s conception of form, see Pip-pin, *Kant’s Theory of Form*.
  35. I discuss the philological side of this issue in more detail in response to Hoke Robinson’s criticism of my earlier accounts in *Idealism and Freedom*, pp. 12–13. For Robinson’s criticism, see “Two Perspectives on Kant’s Appearances and Things in Themselves,” esp. pp. 419–22.
  36. For a discussion of the relevant literature on the topic see S. Morris Engel, “Kant’s Copernican Analogy: A Re-examination,” *Kant-Studien* 59 (1963), pp. 243–51; and especially Norwood Russell Hanson, “Copernicus’ Role in Kant’s Revolution,” *Journal of the History of Ideas* 20 (1959), pp. 274–81. The latter points out that Kant himself nowhere uses the expression “Copernican revolution” to characterize his own thought, and that the explicit comparison of his own procedure to that of Copernicus consists simply in noting that they both tried an alternative hypothesis when existing theories proved unsatisfactory. A more recent and nuanced treatment of the topic, which attempts to trace the connections between Kant’s philosophical views and his cosmological speculations, is provided by Pierre Kerszberg, “Two Senses of Kant’s Copernican Revolution,” *Kant-Studien* 80 (1989), pp. 63–80.
  37. Although Kant does not refer to the synthetic *a priori* at this point, I take it as evident that this is what he had in mind, since the problem he points to does not arise in the case of analytic judgments. Presumably, the reason for this omission is that Kant viewed the analytic-synthetic distinction as a crucial discovery on his part, which he first presents in the Introduction. We shall deal with this distinction in chapter 4, in connection with Kant’s account of judgment.
  38. Bennett, *Locke, Berkeley, Hume*, pp. 136–37.
  39. Berkeley, *The Principles of Human Knowledge*, §6.
  40. *Ibid.*, §3.
  41. *Ibid.*, §132, and *An Essay Towards a New Theory of Vision*, §§79–87. There is a similar analysis in Hume, *A Treatise of Human Nature*, bk. 1, Part 2.
  42. The above analysis should be contrasted with the non-phenomenalist reading of many of the same texts, as well as some similar passages in *On a Discovery*, by Rae Langton (*Kantian Humility*, pp. 140–61 and 186–204). As already noted, Langton interprets Kant as a scientific realist and finds in the *Critique* rudiments of a causal theory of knowledge: empirical objects (composed of forces) are real because they can affect us. Thus, for her everything turns on the existence of a causal relation between the perceiver and the affecting entity, even if the entity (because of the limitations of our sensory apparatus) is not actually perceivable. By contrast, I have emphasized (as I believe Kant clearly does) the law-governedness of the connection between actual perceptions and unperceived (but inferred) entities. Accordingly, on my reading there is no need that something be able actually to affect us in order to count as empirically real or, equivalently, as an object of possible experience. Otherwise, Kant could not talk, for example, about the reality of objects in the distant past, as he certainly intended to do.



43. For a more detailed response to many of these criticisms, see Allison, *Idealism and Freedom*, pp. 3–26.
44. Van Cleve, *Problems from Kant*, pp. 143–50. Van Cleve provides essentially the same analysis in “The Argument from Geometry, Transcendental Idealism, and Kant’s Two Worlds,” in *Minds, Ideas, and Objects*, ed. by Cummins and Zoeller, pp. 296–300.
45. Van Cleve, *Problems from Kant*, p. 147.
46. *Ibid.*
47. *Ibid.*, p. 148.
48. *Ibid.*
49. *Ibid.*, pp. 148–49.
50. Van Cleve characterizes his dualistic interpretation of the thing in itself–appearance distinction as a “qualified two-world” view because it involves construing Kantian appearances as intentional objects or “logical constructions out of states of perceivers,” rather than as entities with a distinct existence. See “The Argument from Geometry,” pp. 295–96 and *Problems from Kant*, p. 142.
51. For reasons that will become clear in connection with the discussion of Ameriks’s objection, the locution ‘may be justifiably claimed to have x’ is intended to replace ‘really has x’.
52. Ameriks, “Kantian Idealism Today,” *History of Philosophy Quarterly* 9 (1992), esp. pp. 334–36. I initially responded to Ameriks’s critique in *Idealism and Freedom*, pp. 17–21. The present response is a significantly revised version of this earlier one.
53. More recently, a similar critique, with special focus on the problem of freedom and the mind-body problem, has been expressed by Rosas, *Kants Idealistische Reduktion*, pp. 117–133.
54. Ameriks, “Kantian Idealism Today,” p. 334.
55. *Ibid.*
56. Obviously, the relation involving sensibility, that is, to things considered as they appear, is *epistemologically* privileged, since it alone can yield cognition.
57. See Carl Posy, “Transcendental Idealism and Causality,” in *Kant on Causality, Freedom, and Objectivity*, ed. by William A. Harper and Ralf Meerbote, p. 38.
58. For my analysis of this issue, see Allison, *Kant’s Theory of Freedom*, pp. 201–49.
59. This is intended to characterize Kant’s critical position prior to the third *Critique*. The situation is complicated, though not radically changed, by the fact that in the latter work Kant introduces a third source of normativity, namely, judgment, to which he assigns an essential mediating role in providing a kind of bridge between the “realms” of freedom and of nature. For my discussion of this complex issue, see Allison, *Kant’s Theory of Taste*, pp. 195–218.
60. The central importance of this conception has been demonstrated by Grier, *Kant’s Doctrine of Transcendental Illusion*. Accordingly, I shall make substantial use of her work in my discussion of it.

### CHAPTER 3

1. The most interesting and informed discussion of the two senses in which Kant construes things in themselves is provided by Rousset, *La doctrine kantienne de l’objectivité*, pp. 167ff.

2. This discussion is based on the philological analysis of the issue by Prauss, *Kant und das Problem der Dinge an sich*, pp. 13–23. Although there are significant differences, some of which will be brought out later in this chapter, my analysis of this topic is deeply indebted to Prauss's.
3. Prauss locates thirty-seven occurrences of the short form and its variants in volumes 3–5 of the Academy Edition (*Kant und das Problem*, p. 14).
4. Prauss locates 258 occurrences of this form in the same volumes (*ibid.*, pp. 14–15).
5. *Ibid.* pp. 42–43.
6. Prauss notes thirteen passages in which the long form is to be found, in some of which Kant uses *ansehen* or *denken* instead of *betrachten* (*ibid.*, p. 20).
7. We shall see in chapter 5 that Kant characterized the Newtonian and Leibnizian conceptions of space and time in this way, that is, as holding that space and time are either properties or relations of things considered as they are in themselves.
8. A firm adherent of this view is Prichard, who assumed that when Kant construes the distinction between appearances and things in themselves to refer to two entities, he requires things in themselves in order to “produce” appearances (*Kant's Theory of Knowledge*, pp. 73–76). It also seems to have been held by Kemp Smith, *A Commentary to Kant's "Critique of Pure Reason,"* esp. pp. 216–18. In more recent literature, it can be found in Rescher, “Noumenal Causality,” in *Kant's Theory of Knowledge*, ed. by L. W. Beck, pp. 175–83; and Moltke S. Gram, “How to Dispense with Things in Themselves (I),” *Ratio* 18 (1976), pp. 1–15.
9. This interpretation is adhered to by Adickes, *Kant und das Ding an Sich*, p. 5; and by Paton, *Kant's Metaphysic of Experience*, esp. vol. 2, pp. 445–46. In the more recent literature, its proponents include Seidl, “Bemerkungen zu Ding an sich und Transzendentalen Gegenstand in Kants *Kritik der reinen Vernunft*,” *Kant-Studien* 63 (1972), pp. 305–14, and Walsh, *Kant's Criticism of Metaphysics*, pp. 162–67.
10. There are numerous places in the *Critique* in which Kant appears to leave room for analytic judgments involving the *concept* of a purely intelligible entity or the intelligible world. See, for example, B 149, A 276 / B 332, A 286 / B 342–43, A 433 / B 461, A 609 / B 637, and A 635 / B 663. The possibility of construing claims about noumena as analytic is noted by Robert Merrihew Adams (“Things in Themselves,” *Philosophy and Phenomenological Research* 57 (1997), p. 811), who nevertheless expresses certain reservations due to the difficulty of understanding and applying the analytic-synthetic distinction. Whatever difficulties there may be in this distinction, however, I think it is clear that Kant thought such claims to be analytic. And, as I shall argue in chapter 4, Kant's understanding of the analytic-synthetic distinction is inseparable from his account of discursivity.
11. Diss 2: 392; 384.
12. A notable exception to this is Adams, “Things in Themselves.” Although he is both sympathetic to two-aspect readings and emphasizes the significance of the ideas of God and the soul, Adams explicitly equates things in themselves with noumena.
13. Kant later makes this explicit when he writes: “Appearances, to the extent that as objects they are thought in accordance with the unity of the categories, are called *phaenomena*” (A 248–49 / B 305).

14. Confining ourselves to the *Critique*, see, for example, B307, A254/B310, A256/B312, A259/B315, B423n.
15. The emphasis on the methodological basis of the conceptions of the noumenon, the thing as it is in itself, and the transcendental object separates the interpretation provided here from the “as if” or fictional interpretations of Vaihinger and more recently of Eva Schaper, “The Kantian Thing-in-Itself as a Philosophical Fiction,” *Philosophical Quarterly* 16 (1966), pp. 233–43.
16. See A253; A358; and R5554 18: 250.
17. I have analyzed this conception of the transcendental object in “Kant’s Concept of the Transcendental Object,” *Kant-Studien* 59 (1968), pp. 165–86.
18. Nevertheless, this is precisely what was done by Kemp Smith, who, following Vaihinger, uses this identification as one of the prongs of his “patchwork” reading of the A-Deduction. See *Commentary*, pp. 204–20.
19. I initially presented this analysis of a third function for the concept of the transcendental object in response to Hoke Robinson’s posing of the “sameness objection” against what he took to be my two-aspect reading of transcendental idealism. (For Robinson’s critique see “Two Perspectives on Kant’s Appearances and Things in Themselves,” pp. 420ff; and for my response see *Idealism and Freedom*, pp. 12–17). Robinson’s objection, which is often directed against two-aspect readings, concerns the coherence of the claim that it is one and the same thing that is to be considered in two ways (or from two points of view). Although it is sometimes assumed that this commits Kant to a highly implausible one-to-one mapping of the phenomenal and noumenal, I take that to be a red herring. First, it is one thing to distinguish between things (taken collectively) as they are for us in virtue of the sensible conditions of human cognition and as they might be for some putative pure understanding, unburdened by such conditions, and quite another to affirm a one-to-one correspondence or isomorphism between the members of the two domains. Second, Kant himself explicitly denies any such isomorphism at A379 and UE 8: 209n. Nevertheless, this still leaves the problem of characterizing that which is being considered from two points of view, and the answer to this question is provided by the transcendental object.
20. See also A277–78/B333–34, A479/B507, A613–14/B641–42.
21. It should be noted that in the Note to the Amphiboly of Concepts of Reflection, which remains unchanged in the second edition, we find a different version of the distinction between a positive and a negative sense of the noumenon or “intelligible object” (A286–88/B342–44). This distinction is between the idea of a thing “thought through pure categories, without any schema of sensibility” (positive sense), which is said to be impossible, and the “objects of a non-sensible intuition” (negative sense). This negative sense is identical to what later becomes the positive sense.
22. Erik Stenius (“On Kant’s Distinction between Phenomena and Noumena,” *Philosophical Essays Dedicated to Gunnar Aspelin on the Occasion of his 65th Birthday*, pp. 231–45) has argued (p. 241) that “the so-called ‘noumena in the negative sense’ ought not to be called ‘noumena’ at all—but rather the transcendental object, which is completely indeterminate.” In this he is perfectly correct. However, as part of his critique of the second

edition version he proceeds to claim that this leads Kant into a contradiction in terms: "For this means that the idea of a 'thing as it is in itself' will be identified with the idea of a thing that is in no way at all. But then the argument of the Transcendental Analytic as well as that of the Transcendental Aesthetic loses its point." What Stenius misses, and what has been stressed here, is the epistemic thrust of Kant's analysis. The idea of a thing as it is in itself is not identified with the idea of a thing that "is in no way at all," but with that of a thing that for us can be nothing more than a mere something = *X*.

23. Lauener, *Hume und Kant*, p. 129.
24. According to Adickes, *Kant und das Ding an sich*, the affection by things in themselves is unambiguously asserted in the following places in the *Critique*: A44 / B61, B72, A190 / B235, A358, A380, A393, and A494 / B522. Of these the last five refer specifically to the transcendental object. Other passages from the *Critique*, such as A288 / B344 and A613–14 / B641–42, could be added to this list, as well as many from other works.
25. See chapter 1, note 20.
26. This point has been emphasized by Gram, who makes it the centerpiece of his interpretation of things in themselves and transcendental idealism. (See *The Transcendental Turn*, pp. 2, 41–50 and passim.) Moreover, this contrasts with standard readings, which tend to assume that affection is simply a species of causality. We have seen, for example, that Langton bases her whole interpretation on an understanding of affection in straightforwardly causal terms (going so far as to make Kant into a proponent of the causal theory of knowledge). Unfortunately, I ignored this point in my first edition.
27. As Gram puts it, "while every case of affection may be accompanied by a case of causation, the two relations are logically independent of each other" (*The Transcendental Turn*, p. 41).
28. The importance of the reference to the transcendental object in the discussions of affection has been noted by others. See especially Herbert Herring, "Das Problem der Affektion bei Kant," *Kant-Studien, Ergänzungshefte* 67 (1953), pp. 65–69. The first part of this important study provides a valuable survey of the literature on the problem of affection. In the second part, Herring develops an interpretation of affection by means of the transcendental object (distinguished from the thing in itself) which differs from the one offered here in its ontological focus. More recently, Henri Lauener (*Hume und Kant*, p. 130) has emphasized both the importance of distinguishing between the thing in itself and the transcendental object and of viewing the latter as "*der rechtmässige Grund (Frage quid iuris) der transzendentalen Affektion, die Notwendigkeit der empirischen verbürgt.*"
29. Jacobi, *David Hume über den Glauben, oder Idealismus und Realismus, Beilage, Ueber den transzendentalen Idealismus*, in *Werke*, ed. by F. Roth and F. Köppen, vol. 2, pp. 291–310.
30. The most forceful proponent of this view is Fichte, especially in his *Zweite Einleitung in die Wissenschaftslehre*, in *Erste und zweite Einleitung in die Wissenschaftslehre*, ed. by Fritz Medicus, pp. 68–75.
31. Vaihinger, *Commentar zu Kants Kritik der Reinen Vernunft*, vol. 2, p. 53.
32. Adickes, *Kants Lehre von der doppelten Affektion unseres Ich als Schlüssel zu seiner Erkenntnistheorie*.
33. See Kemp Smith, *Commentary*, pp. 612ff., Wolff, *Kant's Theory of Mental Activity*, pp. 169ff., 222ff.; Sellars, *Science and Metaphysics*, p. 52.

34. See Lachière-Rey, *L'idéalisme kantien*, pp. 450–63; Bird, *Kant's Theory of Knowledge*, pp. 18–35; Rousset, *La doctrine*, pp. 190–97; Prauss, *Kant und das Problem*, pp. 192ff.; and Gram, *The Transcendental Turn*, pp. 11–39.
35. Prauss, *Kant und das Problem*, pp. 192–207.
36. Ibid.
37. These include A28, A166 / B208, A213 / B260, and Pro 4: 290; 85. For a thorough discussion of the empirical affection passages see Adickes, *Kants Lehre*, pp. 5–15.
38. A similar point is argued for by Gram, *The Transcendental Turn*, pp. 49–52 and passim.
39. Some such line of thought appears to be at work, at least implicitly, in Prauss's analysis.
40. Bird, *Kant's Theory*, p. 69.
41. Precisely the opposite is claimed by Prauss (*Kant und das Problem*, p. 103, n. 22), who asserts that the present passage contains one of the most extreme examples of the “*transzendent-metaphysischen Entgleisungen*” to be found in the Kantian corpus.
42. Kant refers to “the transcendental matter of all objects as things in themselves” (A143 / B182) and to “matter in the transcendental sense,” which is defined as the “determinable in general” (A266 / B322) and equated with the “things themselves which appear” (A268 / B324). For a pre-critical version of this conception, see Diss 2: 389; 379.
43. Eberhard conflated these grounds with the simple parts of which space and time are allegedly composed. In the present context Kant is concerned with underlining the distinction between ground (which is transcendental) and matter or part (which is empirical). He also, however, attacks Eberhard's conception of space and time as composed of simples. For a discussion of this, see Allison, *Kant-Eberhard Controversy*, pp. 117–23.
44. The idea that the thing in itself and associated notions are to be understood in relation to a Kantian metalanguage is suggested by D. L. C. MacLachlan, “The Thing in Itself Appears in a Meta-Language,” *Proceedings of the Eighth International Kant Congress*, vol. 2, pp. 155–61.



#### CHAPTER 4

1. Insofar as this account is connected with the conception of sensibility as providing both a necessary and an *a priori* component to cognition, it constitutes a rejection of rationalist epistemologies as well. But the emphasis on the necessity of the spontaneity of the understanding and the insufficiency (though indispensability) of sensible intuition cuts decisively against empiricism.
2. This is something of an oversimplification, since, as we shall see, Kant, following the tradition, also includes judgment (*Urteilkraft*) as one of the “higher” cognitive faculties (the “lower” being sensibility). But judgment (as a distinct cognitive faculty) plays a relatively modest role in the first *Critique*, and it is only in the third *Critique* that its autonomy is acknowledged, with the recognition of a purely reflective function of judgment that brings with it its own *a priori* principle (the purposiveness of nature) and its peculiar mode of autonomy (heautonomy). For my analysis of this third *Critique* conception of judgment, see *Kant's Theory of Taste*, esp. chapter 1.
3. Elsewhere Kant makes the same point by characterizing the understanding (the faculty of concepts) as the faculty of rules (see A126, A132 / B172). In the secondary literature,

- this feature of Kant's position has been most strongly emphasized by Wolff, *Kant's Theory of Mental Activity*, pp. 70, 130, 142, 220, 323–24.
4. Longuenesse, *Kant and the Capacity to Judge*, pp. 48–50.
  5. It is this sense of 'concept' that Kant has in mind when he refers to a savage who lacks the concept of a house (JL 9: 33; 44–45). I discuss this in *Kant's Theory of Taste*, esp. pp. 26–28.
  6. See also R2876 and R2878 16: 555, 557.
  7. The importance of this problem for Kantian epistemology is emphasized by Longuenesse, *Kant and the Capacity to Judge*. Working on the basis on Longuenesse's analysis, but differing from her in some details, I attempt to deal with it in *Kant's Theory of Taste*, chapter 1. A particular sharp formulation of the problem, which likewise influenced my analysis, is provided by Hannah Ginsborg, "Lawfulness without a Law: Kant on the Free Play of Imagination and Understanding," *Philosophical Topics* 25 (1997), pp. 37–81.
  8. As Longuenesse notes, Kant is here partially anticipated by Locke, who constantly emphasizes the "workmanship of the understanding" in the formation of concepts (*Kant and the Capacity to Judge*, p. 119 and passim). But the anticipation is only partial. I discuss the issue in *Kant's Theory of Taste*, chapter 1, and in more detail in "The Critique of Judgment as a 'True Apology' for Leibniz," in *Kant und die Berliner Aufklärung, Akten des IX. Internationalen Kant-Kongresses*, ed. by V. Gerhardt, R.-P. Horstmann, R. Schumacher, vol. 1, pp. 286–99.
  9. Both criterial features of an intuition are affirmed in the Transcendental Aesthetic, albeit in different places. Thus, immediacy is introduced in the very first sentence, where Kant also states that because of this immediacy all thought is directed to intuition (as the only way in which it links up with its object) (A19 / B 33). Correlatively, we shall see in chapter 5 that singularity is emphasized when, in the Metaphysical Expositions, Kant argues that the representations of space and time are intuitions rather than concepts.
  10. Hintikka, "On Kant's Notion of Intuition (*Anschauung*)," in *The First Critique*, ed. by T. Penelhum and J. MacIntosh, pp. 38–53, and "Kant on the Mathematical Method," *Kant Studies Today*, ed. by L. W. Beck, pp. 117–40. Hintikka's denigration of the immediacy criterion has also been criticized, albeit from a quite different perspective, by Parsons, "Kant's Philosophy of Arithmetic," in *Philosophy, Science and Method*, ed. by S. Morgenbesser, P. Suppes, and M. White, pp. 568–94, esp. pp. 578–80, and by Wilson, "Kant on Intuition," *Philosophical Quarterly* 25 (1975), pp. 247–65, esp. p. 252.
  11. This was already noted by J. S. Beck in his letters to Kant of Nov. 11, 1791 (Br II: 311; 396), and May 31, 1792 (Br II: 338–39; 414). For Kant's response see his letter of July 3, 1792 (Br II: 347–48; 421). It has been brought up again in connection with Hintikka's interpretation of 'intuition' as the equivalent of a singular term by Manley Thompson, "Singular Terms and Intuitions in Kant's Epistemology," *Review of Metaphysics* 26 (1972), pp. 314–43.
  12. de Vleeschauwer, *La déduction transcendentale dans l'oeuvre de Kant*, vol. 2, p. 44, points to the difference between Kant and Aristotle on this issue. For Aristotle the unity of a representation is attributed to sensibility, and it is derived from the ontological unity of the thing. Consequently, sensibility itself yields a representation of an object, which is just what Kant denies.

13. Beck reiterates essentially the same point in a letter to Kant of May 31, 1792 (Br II: 338; 414).
14. Walsh, *Kant's Criticism of Metaphysics*, p. 15.
15. I am indebted to Lewis White Beck for the distinction between these senses of 'intuition', which he suggested to me in his comments on an earlier draft of this chapter. It should be noted that the distinction between indeterminate and determinate intuition concerns only the first of these senses.
16. See also MAN 4: 475–76; 190.
17. This is specifically claimed by de Vleeschauwer, *La déduction transcendentale*, vol. 2, pp. 46–47, 131–34.
18. Here I am in essential agreement with Longuenesse, who discusses the topic in detail. See *Kant and the Capacity to Judge*, pp. 81–90.
19. See Paton, *Kant's Metaphysic of Experience*, vol. 1, p. 251.
20. Since Kant changed '*Anschauungen*' to '*Erscheinungen*' in his own copy of the first edition of the *Critique*, there is a dispute regarding the correct reading of the text. Following Paton (*Kant's Metaphysic of Experience*, vol. 1, p. 253, n. 3), and Raymond Schmidt, I take *Anschauungen* to be the proper reading. It should be noted, however, that given the distinction between three senses of 'intuition', nothing very much rides on this textual question because 'intuition' here must be taken to mean the intuited and this, for Kant, is always an appearance. The essential point, therefore, is not the problematic occurrence of 'intuition' in the passage; it is rather the claim that within judgment a concept is related to a "given representation" that is itself immediately related to an object.
21. Both the difference between Kant's conception of general logic and our contemporary conceptions of logic and the disastrous implications for the interpretation of Kant that arise from the failure to recognize this difference has been emphasized by Longuenesse, *Kant and the Capacity to Judge*, p. 5 and passim. As she points out, Kant understands general logic in the manner of the *Port-Royal Logic* as concerned with the nature of discursive thought and "logical form" as the universal rules of such thinking. We shall return to this issue in connection with the analysis of the Metaphysical Deduction in chapter 6.
22. Paton, *Kant's Metaphysics of Experience*, p. 251, n. 3.
23. Similar accounts are contained in numerous other *Reflexionen*, especially those in the "*Losse Blätter aus dem Duisburgischen Nachlass*," 17: 643–73.
24. The problem will reemerge in chapter 6.
25. In his well-known critique of the ontological argument (A598 / B626), Kant denies that existence is a real predicate or determination because it does not add any content to the description of a thing to say that it exists. He does not, however, deny that it is a logical predicate; consequently, even existential judgments can be said to have two predicates. The point will taken up again in chapter 14.
26. Prauss, *Erscheinung bei Kant*, pp. 86–87. As textual support he cites A760 / B788, where Kant clearly contrasts the two notions. It should also be noted, however, that Kant does on occasion equate them. A case in point is A788 / B816. Despite this verbal inconsistency on Kant's part, Prauss is correct, since Kant's analysis of judgment requires distinguishing between objective validity and truth. This conclusion has also been endorsed by Longuenesse, *Kant and the Capacity to Judge*, p. 82.

27. This issue is dealt with by Rainer Stuhlmann-Laeisz in *Kants Logik*, pp. 28–53. He draws a distinction between acceptable (*verträglich*) and unacceptable (*unverträglich*) judgments and between transcendental and empirical truth. He notes that Kant himself uses the notion of transcendental truth with respect to concepts at A220 / B 268, wherein it is equated with objective reality. Building upon Kant's use, Stuhlman-Laeisz suggests that the transcendental truth of a judgment can be understood as its agreement with the conditions of the possibility of experience, and thus with an object of possible experience. Consequently, an empirical judgment can be acceptable if it has transcendental truth, even if it is false in the empirical sense, that is, does not agree with the actual object to which it makes reference. The judgments of transcendental metaphysics, on this view, are unacceptable, because of this lack of transcendental truth.
28. See Longuenesse, *Kant and the Capacity to Judge*, pp. 86–90. The key Kantian text where this distinction is drawn is R3051 16: 633.
29. See Allison, *The Kant-Eberhard Controversy*, pp. 37–38.
30. Beck, "Can Kant's Synthetic Judgments Be Made Analytic?" in *Kant: Disputed Questions*, ed. by Moltke S. Gram, pp. 228–46, esp. pp. 232–35.
31. This is affirmed in the criticism of what he takes to be the "Kant-Beck thesis" by Moltke S. Gram, "The Crisis of Synthetcity: The Kant-Eberhard Controversy," *Kant-Studien* 71 (1980), pp. 155–80.
32. This is perhaps the most fundamental way in which Kant's account of the analytic-synthetic distinction differs from contemporary accounts and criticisms stemming from the work of Quine. For Quine and those working in his tradition, the focus of the problem lies in the notion of analyticity; whereas for Kant, working in the tradition of Leibniz, it lies in the understanding of a synthetic judgment and, of course, how one is possible *a priori*.
33. Also relevant at this point is the different, but related, objection that was first raised by Eberhard's astute associate J. C. Maass and later developed by C. I. Lewis. According to this line of objection, the distinction between analytic and synthetic judgments is a variable one because any given judgment can be classified as either analytic or synthetic, depending on how one happens to characterize the subject concept. I have dealt with this objection at some length in the Introduction to *The Kant-Eberhard Controversy*.
34. This characterization of analytic judgments as involving a "formal extension" of knowledge requires the distinction of such judgments from tautologies. Unfortunately, Kant is inconsistent on this point. For example, in Fort 20: 322; 404, he draws just such a distinction; though in JL 9: 111; 607, he treats tautologies as a subset of analytic judgments. For a discussion of this issue see Vleeschauwer, *La déduction transcendentale*, vol. 3, p. 406.
35. As noted by L. W. Beck, in "Can Kant's Synthetic Judgments be Made Analytic?" p. 230, Kant himself makes the point in Reflexion 4674 17: 645, when he remarks that in analytic judgments "*das x fällt weg*."
36. UE 8: 230–31; 319–20.
37. Beck, "Can Kant's Synthetic Judgments Be Made Analytic?" p. 231, and "Kant's Theory of Definition," in *Kant: Disputed Questions*, p. 225.
38. See A728 / B 756, where Kant asks, "[W]hat would be the point of defining such a con-



cept? since when, e.g., water and its properties are under discussion, one will not stop at what is intended by the word ‘water’ but rather advances to experiments. . . .” As Beck notes in his comments on the passage, “Description suffices; definition which aims at being more than nominal is a useless presumption” (“Kant’s Theory of Definition,” p. 223). Kant’s point seems to be that judgments involving such empirical concepts are normally not analytic; but if one does explicitly endeavor to make an analytic judgment, that is, appeal to meaning, one can appeal only to a purely nominal definition, “what is intended by the word.” This makes the judgment arbitrary. One is perhaps tempted to say that such judgments about words, in contrast to the intension of concepts, are empirical claims about linguistic usage. Kant, however, does not seem to have considered that possibility.

39. Kant himself limits this claim to theoretical judgments, since he recognizes synthetic judgments of practical reason that do not involve a reference to intuition; for example, “An absolutely good will is one whose maxim can always have as its content itself considered as a universal law.” I discuss this issue in *The Kant-Eberhard Controversy*, p. 74. For a different view on this point see Gram, “Crisis of Synthetcity,” p. 168, n. 24.
40. Although this is obscured by the way in which Kant formulates the analytic-synthetic distinction in the *Critique*, he does clarify matters considerably in his response to Eberhard and in the related correspondence with Reinhold. See especially Br II: 38; 301.
41. I am not here concerned with the question of the cogency of Kant’s philosophy of mathematics, merely with the explication of his claim that synthetic *a priori* judgments require pure intuitions for their grounding.

## CHAPTER 5

1. That in the second edition Kant lists four possibilities, rather than merely the three suggested in the first edition, has been noted by Falkenstein, *Kant’s Intuitionism*, p. 147. As he correctly notes, this was already pointed out by both Vaihinger, *Commentar zu Kants Kritik der reinen Vernunft*, vol. 2, pp. 131–34, and Martin, *Kant’s Metaphysics and Theory of Science*, pp. 11–12. Moreover, there are several other texts in which Kant clearly distinguishes between these possibilities, including Diss 2: 400; 393 and 403; 397, R5298: 18, 146–7, and R5404 18: 174. Nevertheless, as is argued below, at least from the time of the Dissertation, Kant effectively assumes that there are only two alternatives to be refuted: the Newtonian and the Leibnizian positions.
2. For a parallel discussion of time, see Diss 2: 400; 394.
3. See Martin, *Kant’s Metaphysics and Theory of Science*, p. 14. By contrast, Spinoza, who might also be associated with such a view, held that extension is an attribute (not an accident or modification) of God.
4. As Martin points out, Kant tended to identify the substantival view of space with Epicurus (not Newton) and presumably did not take it seriously because it was rejected by both Newton and Leibniz. (See *Kant’s Metaphysics and Theory of Science*, pp. 12–13.) I am not familiar with any thinker to whom one might attribute a substantival view of time.
5. Although in these sections Kant refers to the *concepts* of space and time, I am following the fairly common procedure of substituting the generic term *representations*, on the

grounds that it is an essential part of Kant's project to show that these representations are intuitions *rather than* concepts. Thus, on this reading, his initial use of 'concept' is to be taken in a loose, non-technical sense. For a different view on this point, see Falkenstein, *Kant's Intuitionism*, esp. pp. 63–64.

6. The first alternative is advocated by Vaihinger, *Commentar zu Kants Kritik der reinen Vernunft*, vol. 2, p. 197; the second by Kemp Smith, *A Commentary to Kant's "Critique of Pure Reason"*, pp. 99–105, and Paton, *Kant's Metaphysic of Experience*, vol. 1, pp. 110–14; and the third by Falkenstein, *Kant's Intuitionism*, p. 193 and *passim*.
7. Perhaps the strongest evidence for this reading as an expression of Kant's intent is that he appealed only to the first arguments in the *Inaugural Dissertation*. Admittedly, this does not provide decisive evidence regarding his views in 1781 and 1787; but it does suggest that in 1770 he thought that the first apriority argument sufficed to establish the apriority of space and time in a sense sufficiently strong to warrant conclusions about their subjective status as forms of the sensible world or appearances. Of course, this does not address the cogency of the arguments, so it remains open to the critic to maintain that the first argument in fact accomplishes less than (or something different from) the second. That issue will be dealt with below.
8. This is precisely the mistake I made in the first edition. The error has been pointed out by Falkenstein, *Kant's Intuitionism*, pp. 163–65 and Warren, "Kant and the Apriority of Space," *The Philosophical Review* 107 (1998), esp. 184–92. Accordingly, I have significantly revised my interpretation of the argument to meet their criticisms, while still maintaining what I take to be its main epistemological thrust.
9. This point is emphasized by Warren as part of his critique of my initial interpretation of this argument ("Kant and the Apriority of Space," esp. pp. 187–92).
10. This is suggested by Kant's previously noted discussion of the Fourth Paralogism in the first edition. There, in criticizing empirical idealism, Kant suggests that such a position fails to distinguish between the empirical and the transcendental senses of '*ausser uns*'. The former is defined in spatial terms, and the latter in terms of "pure concepts of the understanding" (A367–71). Since this is equivalent to the distinction implicit in the characterization of outer sense, it suggests that, from Kant's point of view, the empiricist about space and the empirical idealist / transcendental realist may suffer from the same confusion.
11. In commenting on this passage, Warren calls attention to the 'this' [*diese*] qualifying 'outer experience' and correctly notes that it suggests an ambiguity in Kant's claim. It could refer either to outer experience in general or merely to the limited outer experience involving the spatial relations at issue in the argument ("Kant and the Apriority of Space," p. 206, n. 35). We shall return to this issue in connection with the second apriority argument. For the present, it must suffice to note that, even on the qualified reading, the representation of space is still claimed to be a condition of outer experience, which is what Warren apparently denies.
12. *The Leibniz-Clarke Correspondence*, Leibniz's Fifth Paper, §47, p. 69.
13. *Ibid.*, p. 70.
14. See Allison, *Kant-Eberhard Controversy*, pp. 35–36.

15. Much the same line of objection is also raised by Paton (*Kant's Metaphysic*, vol. 1, p. 112).
16. The point is suggested by Warren, "Kant and the Apriority of Space," p. 211. In the light of his analysis of the objection, I have revised my account of the Kantian response from that offered in the first edition.
17. D. P. Dryer, *Kant's Solution for Verification in Metaphysics*, p. 173.
18. The basic point is noted, albeit in somewhat different ways, by both Falkenstein, *Kant's Intuitionism*, pp. 163–65, and Warren, "Kant and the Apriority of Space," p. 210. Once again, I am here revising my analysis in the light of their readings, which are largely correctives to my initial treatment of the argument.
19. As Martin has pointed out (*Kant's Metaphysics and Theory of Science*, pp. 30–34), the schema of this argument is Aristotelian, just as that of the first argument is Platonic. The references to "representation" have been added to reflect the explicitly epistemological thrust of the Kantian version.
20. Kemp Smith, *Commentary*, p. 103.
21. See Julius Ebbinghaus, "Kants Lehre von der Anschauung a priori," in *Kant: Zur Deutung seiner Theorie von Erkennen und Handeln*, ed. by Gerold Prauss, p. 49.
22. For a similar reading, see Dryer, *Kant's Solution*, p. 175.
23. The contrary is maintained by Falkenstein, *Kant's Intuitionism*, pp. 203–10.
24. These features are emphasized by Falkenstein, *Kant's Intuitionism*, *passim*.
25. On this point I am following Warren, "Kant and the Apriority of Space," p. 206, esp. note 35. He points to the difference between the claims of the two arguments and notes that they can be reconciled by adding the above-mentioned claim. This is to be contrasted with the interpretation of the argument by Ted Humphrey ("The Historical and Conceptual Relations between Kant's Metaphysics of Space and Philosophy of Geometry," *Journal of the History of Philosophy* 11 [1973], pp. 503–04). According to Humphrey, this argument, which is not present in the Inaugural Dissertation, is intended to establish a stronger form of necessity than the first. This, he contends, reflects the critical turn in Kant's epistemology between 1770 and 1781. More recently, Falkenstein (*Kant's Intuitionism*, p. 193) has argued for a similar view, suggesting that the first argument is concerned with a relatively weak and negative sense of '*a priori*', where it means "not abstracted from comparison of sensations or the matters of appearance," and the second with a stronger one, where it means "present prior to any constitution of the objects that appear through intellectual processing." Nevertheless, I find such readings unconvincing, since already in the first argument Kant maintains that the representation of space is *a priori* in the sense of being a condition of outer experience.
26. See Parsons, "The Transcendental Aesthetic," in *The Cambridge Companion to Kant*, ed. by Paul Guyer, p. 69.
27. See Vaihinger, *Commentar*, vol. 2, p. 205.
28. The appeal to possible worlds here will not help, for in the sense that we can speak of other possible worlds, we can also speak of other possible spaces.
29. As we shall see when we come to the Antinomies, Kant regards the concept of the world as an idea of reason, but this subtlety may be ignored for present purposes.
30. Admittedly, here Kant does not refer to the immediacy of the representation of space,

- which is one of the problems with the argument. (On this point see Parsons, “The Transcendental Aesthetic,” p. 70). Nevertheless, this seems implicit in Kant’s account of how concepts of space and time are derived. We shall return to this issue in section C, in connection with a consideration of the “givenness” of space.
31. This is cited by Vaihinger, *Commentar*, vol. 2, p. 233, and referred to by Paton, *Kant’s Metaphysic*, vol. 1, p. 122.
  32. The following is based largely on the suggestive account by Kirk Dalles Wilson, “Kant on Intuition,” *Philosophical Quarterly* 25 (1975), pp. 252–56. Much of Wilson’s analysis can be accepted independently of his more controversial claim that Kantian intuitions exhibit a mereological structure. The latter is an interesting suggestion, but one that is not really germane to the present considerations. For a more recent discussion of this topic, which builds on the work of Wilson, see Falkenstein, *Kant’s Intuitionism*, pp. 219–22.
  33. See Buroker, *Space and Congruence*, p. 73.
  34. Kant need not be read as claiming that an intuition necessarily *does* involve infinity, as if this property could be derived from the concept of an intuitive representation. As Rolf P. Horstmann has suggested (“Space as Intuition and Geometry,” pp. 24–25), Kant can also be read as arguing that the representation of space is intuitive, *even if* space is assumed infinite. Since Kant himself regarded the latter as established mathematically, this distinction may not have seemed important to him. Nevertheless, it is of some relevance to the question of the relationship between the Metaphysical Exposition and Kant’s views about geometry.
  35. This reading is even more strongly suggested by Kemp Smith’s rendering as “coexist *ad infinitum*.” But the Guyer and Wood version seems to reflect more accurately the obscurity of the German. Whether or not that is a virtue in a translation is another matter.
  36. This line of objection is developed by Vaihinger, *Commentar*, vol. 2, pp. 257ff. Although the point is often neglected, it is also important to keep in mind here that the Aesthetic is concerned with the infinity of space and time themselves and the Antinomy with that of the *world* in space and time. We shall see that a good deal turns on this difference.
  37. See R6357 18: 682 and the draft of Kant’s letter to Tieftrunk, Br 13: 468.
  38. Surprisingly, in the discussion of the infinity of time, Kant seems to equate unboundedness and limitlessness, though he is elsewhere at pains to distinguish bounds [*Grenzen*] from limits [*Schranken*]. See, for example, Pro 4: 353, 361–62; 106, 114–15. The point is noted by Falkenstein, *Kant’s Intuitionism*, p. 427, n. 14.
  39. See Vaihinger, *Commentar*, vol. 2, pp. 224 ff.; Kemp Smith, *Commentary*, p. 347; Wolff, *Kant’s Theory of Mental Activity*, p. 228.
  40. See LB 17, 20: 22–23. In this note, which is attached to A26 in his own copy of the *Critique*, Kant remarks that the representation of space and time involves the thought of necessity. He further contends, however, that this is not the necessity of a concept (logical necessity) because there is no contradiction involved in the thought of their nonexistence.
  41. Schultz, *Prüfung der kantischen Kritik der reinen Vernunft*, part 2, pp. 41–42.
  42. Melnick, *Kant’s Analogies of Experience*, p. 11.
  43. The phenomenological nature of Kant’s claim is emphasized by Parsons, “The Transcen-

dental Aesthetic,” p. 72. The temporal side of this phenomenology will be explored in chapter 7.

44. Falkenstein, *Kant's Intuitionism*, p. 7.

45. *Ibid.*, p. 8.

46. *Ibid.*, p. 9.

47. Surprisingly, in spite of all the attention he devotes to the issue, Falkenstein dismisses this note on the grounds of its obscurity as not being suitable to play a significant role in an interpretation. (See *Kant's Intuitionism*, pp. 90–91 and also 20–21, 78, 383.) Admittedly, this passage is obscure and has been interpreted in a variety of ways. Nevertheless, Falkenstein is apparently overreacting to the questionable reading of this note by proponents of the view to which he is opposed, namely, that space and time are products of the imagination rather than given.

48. In the secondary literature the distinction between a form of intuiting and a form of intuition, together with the attempt to use this distinction to interpret Kant's own distinction between a form of intuition and a formal intuition, has been presented by Buchdahl, *Metaphysics and the Philosophy of Science*, pp. 579–94, 621; and by Krausser, “The Operational Conception of ‘Reine Anschauung’ (Pure Intuition) in Kant's Theory of Experience and Science,” *Studies in the History and Philosophy of Science* 3 (1972–73), pp. 81–87, and “‘Form of Intuition’ and ‘Formal Intuition’ in Kant's Theory of Experience and Science,” *Studies in the History and Philosophy of Science* 4 (1973–74), pp. 279–87. Both, however, tend to conflate the two distinctions with one another. Thus, the notion of a form of the intuited, which seems necessary to capture the claims that Kant makes about space in the Aesthetic, is simply dropped out.

49. The question of whether space as described in the Transcendental Aesthetic should be regarded as a form of intuition (construed as a capacity or disposition to intuit) or as a formal intuition was first raised by Benno Erdmann in his edition of *Reflexionen Kants zur kritischen Philosophie*, pp. 110–11. He claims that the space that is represented as an infinite given magnitude must refer to the mere form of intuition. Vaihinger, *Commentar*, vol. 2, p. 259, argues to the contrary that the merely potential form of intuition is neither finite nor infinite, and thus that it must refer to the determinate formal intuition. According to my analysis, however, neither is right. Vaihinger is correct in his rejection of Erdmann's analysis, but he is incorrect in jumping to the conclusion that it must be a formal intuition. The problem with both is that they see only two alternatives, when in fact there are three.

50. This is also clear from Pro 4: 322; 75.

51. Buchdahl, *Metaphysics*, pp. 579–82.

52. For a similar view, see Humphrey, “Historical and Conceptual Relations,” pp. 483–512; Rolf P. Horstmann, “Space as Intuition and Geometry,” *Ratio* 18 (1976), pp. 17–30; and Baum, “Kant on Pure Intuition,” in *Minds, Ideas, and Objects*, esp. pp. 312–13.

53. Although Kant likewise provides a Transcendental Exposition of time, arguing that it too is a source of *a priori* knowledge, he does not mount a corresponding argument from this to the ideality of time.

54. For a quite different assessment of the situation, see Michael Friedman (*Kant and the Exact Sciences*, pp. 55–135).

55. See Pro 4: 281–84; 32–36, UE 8: 240, R 5962 18: 403.
56. This is indicated especially clearly by Horstmann, “Space as Intuition and Geometry,” pp. 27–28.
57. This conclusion, which was also drawn in the first edition, has been criticized by both Guyer and Falkenstein. According to Guyer, the main problem with this reading is that it ignores the strong sense of necessity on which Kant insists and which requires the thesis that spatial form be imposed by the mind. Moreover, he claims that if spatial form is imposed by the mind it must have some metric aspect. Thus, on Guyer’s view, Kant is committed to the thesis that Euclidean (or some other geometry) is necessarily true of space (*Kant and the Claims of Knowledge*, pp. 360–61). But his claim, that if spatial form is “imposed” on objects (a rather strange choice of terms for what for Kant is a matter of *receptivity*) it must have a certain metric aspect, is true but irrelevant. For the present point is merely that nothing about the necessarily Euclidean nature of space serves as an essential premise in the Metaphysical Exposition. Falkenstein’s main objection turns on his assumption that the Transcendental Exposition advances a much stronger conception of apriority than the ones delivered by the Metaphysical Exposition and that this stronger sense is required by transcendental idealism (*Kant’s Intuitionism*, pp. 423–24, n. 2). This issue has already been addressed in the discussion of the two apriority arguments.
58. Kant himself clearly expresses the separability of his ideality thesis from his views about geometry when he claims that the former is “not merely an hypothesis to clarify the possibility of synthetic knowledge *a priori*, but a demonstrated truth” (Fort 20: 268; 360). This passage is cited by Horstmann, “Space as Intuition and Geometry,” p. 26.
59. By such counterparts Kant understood objects which are completely similar to one another with respect to their intrinsic properties, but which cannot be contained within the same spatial parameters. These include both geometrical objects, such as spherical triangles, and physical objects, such as left and right hands. In two texts, Pro 4: 285–6; 81–82 and MAN 4: 483–84; 196–97, Kant appealed to the “paradox” of such counterparts in support of the transcendental ideality of space. The most that such counterparts can provide, however, is support for the thesis that the representation of space is an *a priori* intuition, from which the ideality thesis must still be derived. For a contrasting reading emphasizing the importance of incongruent counterparts for Kant’s idealism, see Buroker, *Space and Congruence*.
60. Unfortunately, Kemp Smith here confounds matters for the English reader by translating this clause as a separate sentence, thereby suggesting that Kant is making a new claim rather than merely explicating the previous one.
61. Regrettably, I was myself at least partly captive to this picture in the first edition of this book, which may account for some of the inconsistencies in my analysis.
62. See Paton, *Kant’s Metaphysic*, vol. I, p. 65.
63. This differs radically from Guyer’s reading. Instead of taking these as complementary conclusions mutually entailed by the Expositions, he assumes that Kant intended only the first conclusion to follow from the Expositions, while the second is supposed to follow from the first (See Guyer, *Kant and the Claims of Knowledge*, pp. 355–56). This reading supposedly supports Guyer’s central interpretive thesis that Kant’s idealism is dog-

- matic rather than critical, since it moves from the dogmatic denial that things in themselves are spatial to the claim that space is merely a form of human sensibility. For my discussion of this issue see my review of Guyer's book ( *Journal of Philosophy* 86 (1989), p. 220) and my *Idealism and Freedom*, pp. 22–24.
64. This is particularly emphasized by Paton, *Kant's Metaphysic*, vol. 1, pp. 174–76.
65. Kant does, of course, appeal to such consequences. See, in particular, B 70–72. But these are clearly supplemental considerations, added only in the second edition, and not the main basis for his case against the Newtonian position, which is already formulated in the first edition.
66. Once again, this is a mistake of which I was guilty in the first edition. Here I am particularly indebted to some of the criticisms of that analysis made by Randolph Wojtowicz, "Prolegomena to a Defense of Transcendental Idealism" (unpublished Ph.D. dissertation, University of California, San Diego, 1994).
67. For similar claims, see UE 8: 241; 328, and Fort 20: 266; 359.
68. Paton, *Kant's Metaphysic*, vol. 1, pp. 101–02.
69. In his explicit discussion of the concepts of matter and form in the Amphiboly of Concepts of Reflection (A266 / B 322), Kant defines 'matter' as "the determinable in general" and 'form' as "its determination." These are intended as perfectly general, or "transcendental," definitions, which apply to both judgments and entities. With regard to the latter, however, it should be noted that a thing's determinations make it into the kind of thing that it is. Consequently, the definition is at least compatible with the implicit understanding of 'form' as condition, which we find in both the Transcendental Aesthetic and the Analytic.
70. The history of the objection is sketched by Vaihinger, who is himself very sympathetic to it; *Commentar*, vol. 2, pp. 134–51. As initially developed by Kant's contemporaries, it took a "strong" and a "weak" form, corresponding to the realistic and the idealistic interpretations of the Leibnizian monadology, respectively. According to the "strong" form, it is deemed possible that space is a form of human apprehension and that things in themselves actually are in space or spatial. According to the "weak" form, it is deemed possible that space is such a subjective form but that the realm of things in themselves (conceived as Leibnizian monads) contains an analogue of space. The objection was reformulated (without these Leibnizian overtones) in the nineteenth century by Adolf Trendelenburg, who wrote: "Even if we concede the argument that space and time are demonstrated to be subjective conditions which, in us, precede perceptions and experience, there is still no word of proof to show that they cannot at the same time be objective forms" (*Logische Untersuchungen*, p. 184). This led to a lengthy and bitter controversy with Kuno Fischer, who attempted to defend Kant. I have dealt with this issue in "The Non-spatiality of Things in Themselves for Kant," *Journal of the History of Philosophy* 14 (1976), pp. 313–21. The current analysis, however, differs considerably from the one offered there.
71. Kemp Smith, *Commentary*, p. 113.
72. Reinhold, *Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens*, pp. 244–47.
73. Vaihinger, *Commentar*, vol. 2, p. 313.
74. For example, in speculating about a supposed object of a non-sensible intuition, Kant re-

marks that “its duration is not a time” (B149), which suggests that he might not rule out noumenal analogs to our sensible forms.

75. Although he does not emphasize the latter point, this is basically the objection raised against the version of my response to the problem offered in the first edition of this book by Falkenstein, “Kant’s Argument for the Non-Spatiotemporality of Things in Themselves,” *Kant-Studien* 80 (1989), pp. 265–83. Fortunately, as we are about to see, he also suggests the response to this objection, which I gratefully adopt as a more or less friendly amendment, even though he himself apparently does not see it in quite this way. For my initial use of Falkenstein’s analysis, see *Idealism and Freedom*, pp. 10–11.
76. Falkenstein, “Kant’s Argument,” esp. pp. 275–82; *Kant’s Intuitionism*, pp. 301–04.
77. Falkenstein, *Kant’s Intuitionism*, p. 305. For Falkenstein’s rejection of the idealistic implication see “Kant’s Argument,” pp. 382–83. It must also be kept in mind that a sharp distinction between Kant’s arguments in the Aesthetic and transcendental idealism is a central theme of *Kant’s Intuitionism*. In other words, Falkenstein advocates a version of the separability thesis.
78. Falkenstein, “Kant’s Argument,” p. 282.
79. *Ibid.*, p. 305. As his note 23 on p. 428 indicates, Falkenstein’s worry about the issue of non-spatiality being treated as a matter of definition seems directed at my own earlier analysis of the issue.

## CHAPTER 6

1. Since Kant himself never explicitly tells us, there is some dispute regarding the precise location of the Metaphysical Deduction. In particular, it is often thought to be contained entirely in §10, which is the third section of the “Clue.” See, for example, Rolf P. Horstmann, “The Metaphysical Deduction in Kant’s Critique of Pure Reason,” *Philosophical Forum* 13 (1981), pp. 32–47. On the reading advocated here, the entire “Clue” may be seen as a metaphysical deduction in a broad sense, since it provides the analysis of judgment and its functions that is essential for the determination of the categories. Nevertheless, §10 will here be taken as the “metaphysical deduction proper,” since it is here that Kant argues for the correspondence between the functions of judgment and the categories.
2. This criticism, which has been reiterated countless times, was sharply expressed by Hegel. For a discussion of this see, Reich, *The Completeness of Kant’s Table of Judgments*, pp. 1–2.
3. Strawson, *The Bounds of Sense*, pp. 81–82.
4. These include Brandt, *The Table of Judgments: Critique of Pure Reason A 67–76; B 92–101*, trans. by Eric Watkins, Wolff, *Die Vollständigkeit der kantischen Urteilstafel*; and Longuenesse, *Kant and the Capacity to Judge*. All these authors are responding to Klaus Reich’s pioneering, yet exceedingly controversial, treatment of the topic (see note 2).
5. This is argued by L. Krüger, “Wollte Kant die Vollständigkeit seiner Urteilstafel beweisen?” *Kant-Studien* 59 (1968), pp. 333–55. In the first edition of this work I was in basic agreement with Krüger’s position.
6. See Pro 4: 318; III, and Kant’s letter to Marcus Herz of May 26, 1789, B II: 51; 313–14.



7. This point is emphasized by Wolff (*Die Vollständigkeit*, p. 180), who criticizes Krüger. (See note 5).
8. On this point I am, again, in agreement with Wolff; though I question his suggestion that the discursivity that is here “dogmatically” assumed is somehow vindicated in the Transcendental Deduction (see *Vollständigkeit*, pp. 177 and 181). I shall return to this issue in chapter 7.
9. See, for example, Axiv, Bxxii, A80–81 / B106–07, and Pro 4: 322–24; 114–16.
10. On the terminological issue, see Wolff, *Die Vollständigkeit*, p. 85, n. 84.
11. This important point is emphasized by both Wolff, *Die Vollständigkeit*, pp. 85–86, and by Brandt, *The Table of Judgments*, pp. 64–65.
12. This point is emphasized by Wolff, *Die Vollständigkeit*, pp. 40–41.
13. What follows is essentially my distillation of the analyses provided by Brandt, Wolff, and Longuenesse. Although much of the terminology and to some extent the way of framing the issue is my own, I am greatly indebted to their work on virtually every significant point.
14. On the peculiarity of modality as a function type, see Brandt, *The Table of Judgments*, pp. 6, 62–63; Wolff, *Die Vollständigkeit*, pp. 124–29; and Longuenesse, “The Divisions of Transcendental Logic and the Leading Thread,” p. 147.
15. As Wolff puts it, for Kant modality as a logical form of judgment is always *de dicto* (*Die Vollständigkeit*, p. 126). This does not mean that Kant rejects all *de re* modality, merely that it has no place in an analysis of judgment from the standpoint of general logic.
16. Kneale, *The Development of Logic*, pp. 355–56.
17. See Brandt, *The Table of Judgments*, esp. pp. 72–84.
18. See Wolff, *Die Vollständigkeit*, p. 161, for the critique of Brandt. It should also be kept in mind, however, that both Brandt and Wolff are responding to Klaus Reich’s treatment of the issue (*The Completeness of Kant’s Table of Judgments*, esp. pp. 101–09).
19. Typical here is the reaction of Kemp Smith, *Commentary*, p. 192.
20. The functions of affirmative and negative judgments according to Kant are, respectively, to extend knowledge and ward off error (See A709 / B737). This limitative function turns out to be of particular importance in Kant’s analysis of complete determination or individuation treated in “The Ideal of Pure Reason,” esp. A571–80 / B579–608.
21. The problem is complicated by the fact that the introduction of “relation” as a title was an innovation on Kant’s part, replacing the standard division between simple and complex judgments, where the latter would include, in addition to hypothetical and disjunctive judgments, also copulative judgments and perhaps other forms as well. On this point see Kemp Smith (*Commentary*, pp. 192–93), who predictably asserts that Kant’s omission of this form is further evidence of the illicit nature of his procedure in selecting only judgment forms that will yield the desired categories.
22. The parenthetical additions are my emendations of Kant’s characterization, but they seem to be required by the conception of a copulative judgment.
23. They are discussed by Georg Friedrich Meier, *Auszug aus der Vernunftlehre*, §304, which Kant used as the textbook for his logic lectures. Meier’s text is reprinted in KGS, vol. 16, which contains Kant’s *Reflexionen* dealing with logic.
24. See also R3III 16: 663.

25. This again points to the gulf separating Kant's analysis of judgment from the truth-functional approach of contemporary logic, where conjunction plays a significant role precisely because the elements linked by logical connectives have independent truth values.
26. The point is raised by Kemp Smith, *Commentary*, pp. 193–94.
27. Although I take a somewhat different and more direct approach, my analysis of this note is heavily indebted to Wolff's detailed and systematic discussion. See *Die Vollständigkeit*, pp. 16–74.
28. Kant reiterates this point in a letter to J. S. Beck, who had suggested that *all* synthetic divisions must be trichotomous. See Br 11: 394; 445.
29. In R3030 16: 622–23 and R5854 18: 369–70, Kant connects trichotomous divisions with the unity of consciousness.
30. I am here following Wolff, *Die Vollständigkeit*, pp. 172–73.
31. Longuenesse, *Kant and the Capacity to Judge*, p. 5.
32. *Ibid.*, pp. 5–6.
33. *Ibid.*, p. 5.
34. The issue of psychologism is noted by Longuenesse, *Kant and the Capacity to Judge*, pp. 6–7.
35. Strawson, *The Bounds of Sense*, p. 32.
36. A prime example of this approach is Kitcher, *Kant's Transcendental Psychology*. For my critique of Kitcher, see *Idealism and Freedom*, pp. 53–66.
37. Longuenesse, "The Divisions of the Transcendental Logic and the Leading Thread," p. 143, and *Kant and the Capacity Judge*, pp. 3–6 and *passim*.
38. Insofar as Kant understands a function as a unifying rule, he is clearly construing the term in its mathematical sense as the law underlying an operation. (This sense of the term is emphasized by Reich, *The Completeness of Kant's Table of Judgments*, p. 27.) As we have already seen, however, Kant usually understands the term in the Aristotelian or physiological sense.
39. In a first-edition passage, which is also contained in the second, Kant defines this concept as the relation of subsistence and inherence (A80 / B 106). Both of these definitions are merely nominal, however, since Kant denies the possibility of a real definition of any of the pure concepts. On this latter point, see A240–41 / B 300–01.
40. This is emphasized by Bennett, *Kant's Analytic*, p. 183.
41. As we shall see in chapter 12, Kant explores the illicit projection of the ontological concept of substance in connection with the soul or self in the Paralogisms.
42. This is pointed out by Melnick in *Kant's Analogies*, p. 39.
43. Melnick (*Kant's Analogies*, p. 51), suggests the appropriateness of the Rylean notion in this context.
44. Guyer, *Kant and the Claims of Knowledge*, p. 99. In the continuation of the same sentence he makes a similar point about the connection between disjunctive judgments and interaction.
45. The issue is discussed in detail in connection with the Third Analogy by Longuenesse, *Kant and the Capacity to Judge*, pp. 375–87.
46. This goes back at least to the contrast drawn in the Inaugural Dissertation between the logical and the real use of the intellect (*intellectus*), where Kant denied that the intellect

- has a real use in connection with sensible cognition (Diss 2: 394; 386). In the *Critique*, where the main concern of the Transcendental Analytic may be described as demonstrating that the understanding has a real use, albeit one limited to possible experience, he does not use that expression. Nevertheless, the contrast between the logical and real uses of the understanding is certainly implicit in the entire account, since the first section of the “Clue” is given the heading: “On the logical use of the understanding in general” (A67 / B92). Moreover, in the Introduction to the Dialectic Kant distinguishes between the logical and the “pure” (real or metaphysical) use of reason in a way that suggests that the same distinction is applicable to the understanding (A303–09 / B359–66).
47. Following the suggestion of de Vleeschauwer and Ian Proops (in an unpublished paper), I have argued in *Kant's Theory of Taste* (pp. 67–84 and passim) that the Metaphysical and Transcendental Deductions are concerned, respectively, with the *quid facti* and the *quid juris*. Nothing in the present discussion, however, turns on this.
  48. See Longuenesse, “The Divisions of the Transcendental Logic and the Leading Thread,” p. 149.
  49. In the older literature, this was emphasized by Reich and following him, Paton (*Kant's Metaphysic*, vol. 1, pp. 281–302).
  50. See Kemp Smith, *Commentary*, pp. 178–80, and Wolff, *Kant's Theory of Mental Activity*, pp. 68–77.
  51. See Reich, *The Completeness*, pp. 8–10.
  52. See Paton, *Kant's Metaphysic*, vol. 1, p. 288.
  53. *Ibid.*, p. 290.
  54. Listed in chronological order, see also: Pro 4: 324; 116, MAN 4: 474; 189 B128, and Fort 20: 272; 363.

## CHAPTER 7

1. It is not, however, completely obscured. A case in point is the suggestion that the purpose of such a deduction is “to make comprehensible this relation of the understanding to sensibility and by means of the latter to all objects of experience” (A128).
2. J. Claude Evans has appropriately characterized this worry as the transcendental analogue to Descartes's specter of the evil genius (“Two-Steps-in-One-Proof: The Structure of the Transcendental Deduction of the Categories,” *Journal of the History of Philosophy* 28 [1990], pp. 553–70). As Evans points out, however, it is important to keep in mind the differences between the two specters.
3. Henrich, “The Proof-Structure of Kant's Transcendental Deduction,” in *Kant on Pure Reason*, ed. by Ralph C. S. Walker, pp. 67–68.
4. *Ibid.*, p. 70.
5. See, for example, Raymond Brouillet, “Dieter Henrich et ‘The Proof-Structure of Kant's Transcendental Deduction’: Réflexions Critiques,” *Dialogue* 14 (1975), pp. 639–48; Hans Wagner, “Der Argumentationsgang in Kants Deduktion der Kategorien,” *Kant-Studien*, 71 (1980), pp. 352–66; and Hoke Robinson, “Intuition and Manifold in the Transcendental Deduction,” *Southern Journal of Philosophy* 23 (1984), pp. 403–12.
6. Dieter Henrich, “Diskussion: Beweisstruktur der transzendentalen Deduktion,” in *Prob-*

- leme der "Kritik der reinen Vernunft,"* ed. by Burkhard Tuschling (Berlin: Walter de Gruyter, 1984), pp. 34–96, esp. pp. 41–42. The whole controversy is discussed by Evans, "Two-Steps-in-One-Proof," pp. 554–60.
7. Evans, "Two-Steps-in-One-Proof," pp. 558–60, attempts to defend Henrich's revised version on this score by emphasizing that the "restriction" applies not to intuitions insofar as they already possess unity (independently of the activity of the understanding), but rather to them insofar as they are unifiable by the understanding. On this reading, then, the task of the second part of the Deduction is to show that all human intuition is so unifiable. Although I have nothing to object to in this as an account of the proof structure of the argument, I believe it is somewhat misleading to depict this as a restriction that is removed in the second part. For by depicting it as a restriction on the "range" of intuitions to which the categories apply, Henrich clearly implies that the first part shows (or at least attempts to show) that the categories do apply to a certain range of intuitions. And it is just this that I find questionable.
  8. Hoke Robinson, "Intuition and Manifold in the Transcendental Deduction," pp. 403–12.
  9. Thus, the heading of the very section in which Henrich claims to find the narrower or restricted claim states: "All sensible intuitions stand under the categories, as conditions under which alone the manifold can come together in one consciousness" (B143).
  10. To my knowledge, the problem was first posed, though not resolved, by Rudolf Zocher, "Kants Transzendente Deduktion der Kategorien," *Zeitschrift für philosophische Forschung* 8 (1954), pp. 163–94.
  11. In the first edition of this book I organized the analysis of the two parts of the Deduction around the distinctions between objective validity and objective reality, on the one hand, and *Object* and *Gegenstand*, on the other. The basic idea was that the first part is to be seen as establishing the objective validity of the categories, understood as their validity with respect to objects in a very broad, logical sense (*Objecte*) and the second their objective reality, understood as their applicability to the spatiotemporal objects of human experience (*Gegenstände*). This has been objected to on philological grounds by a number of critics, and as a result of further reflection on the matter I have come to realize that introducing these considerations was both misleading and unnecessary to my main goal of determining the relation between the two parts of the argument. The basic point is simply that the first part of the B-Deduction is concerned with the role of the categories in the *thought* of objects, that is, in judgment, whereas the second is concerned with their role in perception and experience.
  12. A similar analysis is provided by Thöle, *Kant und das Problem der Gesetzmässigkeit der Natur*, esp. pp. 264–69.
  13. Here, as elsewhere, I am modifying the Guyer and Wood translation by adding "to ourselves" in order to capture what I take to be the extremely important reflexive feature of Kant's claim, which is frequently (though not always) lost in their translation.
  14. This objection has been raised in numerous forms, most emphatically by Guyer, *Kant and the Claims of Knowledge*, pp. 140–42. For an interesting response to Guyer on this point, see Keller, *Kant and the Demands of Self-consciousness*, pp. 67–69.
  15. Elsewhere Kant refers to this I as the "I of reflection" in contrast to the "I of apprehen-

- sion" (Anthro 7: 141–42; 22), and the "logical I" in contrast to the "psychological I" (Fort 20: 270).
16. See, for example, Kant's letter to Marcus Herz of May 26, 1789, Br 11: 52; 314, and Anthro 7: 135–36; 16–17.
  17. The latter point has been challenged by Keller, *Kant and the Demands of Self-Consciousness*, p. 68. But since he does acknowledge that Kant allows for the logical possibility of representations *in me* that are nothing for me, I cannot see the point of his disagreement. If representations are in me, then surely there is a sense in which they are mine (as opposed to being someone else's). A case in point is Guyer's example of a dream that I fail to remember upon awakening. My failure to remember does not prevent the dream from being mine, and the dream does not first become such when or if I recall it.
  18. See particularly A 99. In his discussion of the necessity of a threefold synthesis, Kant presents the problem as one of determining the conditions of the representation of a temporal manifold as such. But in the more abstract formulation of the B-Deduction, where the first part explicitly abstracts from the specific nature of human sensibility, the concern is with the conditions of the representation of a sensible manifold as such, that is, simply as a manifold.
  19. The comparison between James and Kant on this point is noted by Kemp Smith, *A Commentary to Kant's "Critique of Pure Reason,"* p. 459, and by Wolff, *Kant's Theory*, p. 106.
  20. Guyer, "Kant's Tactics," p. 184, and *Kant and the Claims of Knowledge*, pp. 134, 136–37, 139–40.
  21. There is an interesting contrast here with the first edition version of the Paralogism, where Kant might be taken to be denying that the principle is analytic, since he denies the possibility of proving "from concepts" the proposition that "many representations have to be contained in the absolute unity of the thinking subject in order to constitute one thought" (A 352). A consideration of the context, however, indicates that the emphasis must be placed on "absolute," which generally has a metaphysical sense for Kant. Moreover, the only thing that he explicitly denies to be analytic is the quite different and manifestly metaphysical proposition "A thought can only be the effect of the absolute unity of the thinking being" (A 353). The original proposition, presumably stripped of its metaphysical sense, is explained as an expression of a necessary condition for apperception. For a different interpretation of this passage see Kemp Smith, *Commentary*, p. 479.
  22. In contrast to the reading offered here, Edwinn McCann ("Skepticism and Kant's B-Deduction," *History of Philosophy Quarterly* 2 [1985], p. 74), and Longuenesse (*Kant and the Capacity to Judge*, p. 67, n. 13) both claim that the principle is based on an analysis of "my representation" or of the notion of a representation being *mine*. I believe, however, that this "mineness" must itself be understood in connection with the concept of discursive thought.
  23. Strawson, *The Bounds of Sense*, p. 32.
  24. For a representative of the latter point of view, see Kitcher, *Kant's Transcendental Psychology*, p. 186. Although he does not accept Kitcher's reading of the Deduction, Keller endorses her rejection of the analyticity thesis (*Kant and the Demands of Self-Consciousness*, pp. 90–91).

25. See also B135.
26. The dogmatic nature of Kant's claim is emphasized by Kemp Smith, *Commentary*, p. 284, and by A. C. Ewing, *A Short Commentary on Kant's "Critique of Pure Reason,"* p. 115, who makes specific reference to the second edition. The latter point is affirmed by Henrich, *Identität und Objektivität*, pp. 7, 21.
27. See also A78 / B103, where Kant initially characterizes synthesis as a "mere effect of the imagination" and describes the latter as "a blind though indispensable function of the soul [or understanding], without which we would have no cognition at all, but of which we are seldom even conscious."
28. See, for example, Kitcher, *Kant's Transcendental Psychology*, pp. 111, 126–27, where she dismisses the idea of a consciousness of synthesis as "synthesis watching." I criticize Kitcher's account in "On Naturalizing Kant's Transcendental Psychology," in *Idealism and Freedom*, pp. 53–66.
29. This point is related to the important critique by Ameriks and Dieter Sturma of Henrich's "reflection theory," which, following what he terms "Fichte's original insight," Henrich attributes to Kant. Very roughly, this theory holds that self-consciousness is to be explained as the result of a subject making an object of itself, and Henrich, following Fichte, claims that any such account is viciously circular. (See Henrich, "Fichtes ursprüngliche Einsicht," in *Subjectivität und Metaphysik*, pp. 188–232). What Sturma and Ameriks point out against Henrich is not just that Kant was not committed to such a theory but, more interestingly, that he anticipated the germ of Fichte's insight, namely, that consciousness is "primitively self-referential." (See Sturma, *Kant über Selbstbewusstsein*, and Ameriks, "Kant, Fichte, and Apperception," pp. 244–49). As Ameriks succinctly puts it, "'I think that x' . . . is *already* a kind of self-consciousness—since it directly expresses a thinking self's activity—even if it is precisely not a reflection upon a distinct 'object-self'" (op. cit., p. 249).
30. This interpretation is suggested by the analyses of Reich, *The Completeness of Kant's Table of Judgments*, p. 34, and of Stuhlmann-Laeisz, *Kants Logik*, pp. 81–83. More recently, a similar analysis has been offered by Wolfgang Carl, "Apperception and Spontaneity," *International Journal of Philosophical Studies* 5 (1997), pp. 147–63.
31. I also do not believe that it should be taken as a bit of proto-cognitive science, as has been averred by a number of commentators, including Kitcher and most recently Andrew Brook, *Kant and the Mind*, p. 4 and passim. Needless to say, I am not denying that Kant's account of apperception cum synthesis is of *any* potential significance to cognitive science. My point is rather that the connection is more complex and indirect than philosophers like Brook and Kitcher, who treat Kant's central claims as broadly empirical and argue for a straightforward appropriation, suggest. For if Kant's account is anything, it is *a priori*.
32. Previously Kant had defined the understanding as the faculty of cognition through concepts (A68 / B93) and as a "faculty for judging" (*Vermögen zu urteilen*) (A69 / B94). In the A-Deduction, he surveys his various definitions of the understanding and suggests that they come together in its characterization as the "faculty of rules" (A126).
33. Nevertheless, there are significant differences between the two accounts, which perhaps

explains why Kant dropped references to an “object in general,” also characterized as a “transcendental object = x,” in the B-Deduction. For when Kant appealed to the latter in the A-Deduction it was to characterize the thought of an object corresponding to and distinct from our representations (See A104–05, 108–09). By contrast, in the B-Deduction, the focus is not on an object considered as distinct from our representations but rather on one considered in abstraction from its manner of givenness. On this point see Zöller, *Theoretische Gegenstandsbeziehung bei Kant*, pp. 148–49.

34. Among those who criticize Kant’s argument on roughly these grounds are Hossenfelder, *Kants Konstitutions-theorie*, pp. 128–30, and Guyer, *Kant and the Claims of Knowledge*, p. 117. A similar worry is also expressed by Carl, “Die B-Deduktion,” in *Kritik der reinen Vernunft*, pp. 197–98, but he finds its resolution in the specification of the unity of consciousness as an objective unity in §18.
35. I discuss this conception of the spontaneity of thought in judgment, understood as an act of “taking as,” and its connection with self-consciousness in more detail in *Kant’s Theory of Freedom*, pp. 36–38, and, in connection with my critique of Kitcher, in *Idealism and Freedom*, pp. 57–64. A similar analysis, which has influenced my own treatment of this topic, is provided by Robert Pippin, “Kant on the Spontaneity of the Mind,” *Canadian Journal of Philosophy* 17 (1987), pp. 449–75.
36. See Longuenesse, *Kant and the Capacity to Judge*, pp. 33–34.
37. In this note Kant is addressing Johann Schultz’s review of Heinrich Ulrich’s *Institutionis logicae et metaphysicae* (1785). A somewhat eclectic Wolffian, Ulrich attempted to reconcile the main positive results of the *Critique* (e.g., the ideality of space and time and the validity of the Principles) with a Leibnizian metaphysics by arguing that, given its own premises, the *Critique* should not limit cognition to objects of possible experience. In his review, Schultz (a close associate of Kant and author of the first commentary on the *Critique*) expresses agreement with Ulrich on several points, while also chastising him for neglecting the Transcendental Deduction. At the same time, however, he partially excuses Ulrich for this neglect on the grounds of its excessive obscurity. In particular, looking back on the A-Deduction from the vantage point of the *Prolegomena*, he notes an ambiguity in Kant’s conception of experience that threatens the entire argument. If, Schultz argues, Kant understands by this judgments of experience, then the claim that experience in that sense presupposes the categories is trivially true. But if he means judgments of perception, then, on the one hand, there is a contradiction between the *Critique* and the *Prolegomena*, and, on the other, the argument fails to establish the necessity of the categories. The importance of these criticisms, coming from someone so close to home, for understanding Kant’s revised strategy in the B-Deduction has been convincingly demonstrated by Thöle, *Kant und das Problem der Gesetzmässigkeit der Natur*, pp. 274–82.
38. In addition to referring to the “logical connection of perceptions in a thinking subject” (see also Pro 4: 304; 98), which is presumably a connection according to the logical functions, Kant also refers to the “logical moments in all judgments” and states that if the very same moments serve as concepts, “they are concepts of the necessary unification of these representations in a consciousness, and so are principles of objectively valid judg-

- ments" (Pro 4: 305; 98). The latter suggests that there are judgments in which these moments *do not* serve as such concepts and which therefore lack objective validity.
39. As Longuenesse points out (*Kant and the Capacity to Judge*, p. 179, n. 26), it is appropriate to translate 'Wenn' here as 'if' rather than 'when' in order to capture the hypothetical form of the judgment. This makes it clear that the change from a judgment of perception to a judgment of experience does not involve a change in logical form.
  40. A discussion of the science behind this example, based on an account by Michael Friedman, is provided by Longuenesse, *Kant and the Capacity to Judge*, pp. 179–80, n. 27.
  41. A similar approach is taken by Rhoda Kotzin and Jörg Baumgärtner, "Sensations and Judgments of Perceptions, Diagnosis and Rehabilitation of Some of Kant's Misleading Examples," *Kant-Studien* 81 (1990), pp. 401–12.
  42. As Kant himself was keenly aware (see Pro 4: 305n; 99), the difficulty with this analysis lies in the notion of necessity, particularly in connection with *a posteriori* judgments. Moreover, similar considerations apply, *mutatis mutandis*, to the parallel discussion in §19 of the *Critique*, where Kant refers to the necessity connected with the empirical judgment: "All bodies are heavy" (B142). The necessity supposedly involved in such judgments can neither be logical, since that would effectively make them analytic, nor of the modal category, since that category does not apply to all judgments and its presence is a criterion of a judgment's apriority, not of its objectivity. In my view, the sense of necessity that pertains even to empirical judgments is a normative one that is grounded in a transcendental condition (the necessary unity of apperception). The basic idea is that the claim that a certain property or relation holds of an object involves an implicit demand for the agreement of others. And underlying this demand is the claim that one is judging the object as it *ought* to be judged. Although the judgment is cognitive and, as such, based on a concept that serves as a rule, this normative necessity is similar to the one involved in a judgment of taste. For my analysis of the latter, see Allison, *Kant's Theory of Taste*, pp. 146–47. For a somewhat different analysis of the issue, which links the necessity of empirical judgments to their derivability from empirical laws, see Longuenesse, *Kant and the Capacity to Judge*, pp. 180–83.
  43. Kant explicitly treats judgments of perception in this way in his treatment of the topic in various versions of his logic lectures. See, for example, LD-W 24: 767; 499 and JL 9: 113; 608. For a discussion of these texts and their relation to the distinction in the *Prolegomena*, see Longuenesse, *Kant and the Capacity to Judge*, pp. 188–95. Moreover, Kant's *Nachlass* provides ample evidence of a complex theory of preliminary judgments dealing with semblances, which was presumably influenced by Lambert's phenomenology, of which judgments of perception may be viewed as a kind. The topic of preliminary judgments and their connection with Kant's theory of empirical knowledge is discussed from a hermeneutical point of view by Svendsen, *Kant's Critical Hermeneutics*, pp. 285–303.
  44. This is likewise suggested by Jürg Freudiger, "Zum Problem der Wahrnehmungsurteile in Kants theoretische Philosophie," *Kant-Studien* 82 (1991), pp. 414–35. Freidiger, however, insists that the B-Deduction actually presupposes judgments of perception (rather than merely perceptions) and that these differ from judgments of experience only in not applying the *schematized* categories. But all that Freidiger succeeds in doing is to indi-



cate the importance of the move from perception to experience for the B-Deduction, a point on which I am in full agreement. Moreover, I find the suggestion that judgments of perception (not to mention perception itself) might involve *unschematized* categories puzzling, since the latter, *ex hypothesi*, involve no reference to sensibility.

45. Here is the major point of my disagreement with Longuenesse on this issue (*Kant and the Capacity to Judge*, pp. 167–97). For she regards it as a virtue rather than a fault of Kant's discussion of judgments of perception in the *Prolegomena* that it does not involve reference to the categories. Although the issues are too complex to pursue here, at bottom this is because she takes this to support her central interpretive thesis about the priority of the logical functions to the categories, particularly when the latter are understood as fully reflected concepts applied in judgments. In my view, however, Longuenesse fails to account for the role of the categories (as distinct from the logical functions) in either part of the B-Deduction. On this latter point, see Allison, "Where Have All the Categories Gone? Reflections on Longuenesse's Reading of Kant's Transcendental Deduction," *Inquiry* 43 (2000), pp. 67–80.
46. Accordingly, I cannot accept the view of Pierre Keller, who suggests that in the B-Deduction Kant restricts judgments to what the *Prolegomena* regarded as judgments of experience (*Kant and the Demands of Self-Consciousness*, p. 85). First, in §19 Kant is clearly giving a generic account of judgment (or of judgment as such), not simply an account of empirical judgments. Second, the distinction between judgments of perception and judgments of experience (albeit in a somewhat different form) is found in texts post-dating the B-Deduction.
47. Kant's fullest account of empirical apperception and its contrast with the transcendental or pure variety is in the *Anthropology*. There Kant first characterizes the latter as the "self-consciousness of reflection" (or understanding) and the latter the "self-consciousness of apprehension" (*Anthro* 7: 134n.; 15). And somewhat later he contrasts "the 'I' of reflection," which contains no manifold, and "the 'I' of apprehension," which does (*Anthro* 7: 141–42; 22). Elsewhere, Kant distinguishes these as the "logical I" and the "psychological I." For a useful account of Kant's conception of empirical apperception and its various formulations, see Carl, *Die Transzendente Deduktion der Kategorien*, pp. 62–65.
48. Keller claims that my account of subjective unity in the first edition of this work failed to account for empirical apperception as a non-objective form of self-consciousness because of my supposed tendency to collapse all self-consciousness for Kant into a consciousness of objective states of affairs' (See *Kant and the Demands of Self-Consciousness* (pp. 82–84 and 242, n. 3). Although I do not believe that the charge that I "collapse" all self-consciousness in the manner he describes is valid, I acknowledge the inadequacy of my previous discussion of empirical apperception, for which the present treatment is intended as a correction. This does not, however, amount to an endorsement of Keller's broader thesis that the goal of the second part of the B-Deduction is to ground self-knowledge (*loc. cit.*, pp. 92–94).
49. See Longuenesse, *Kant and the Capacity to Judge*, pp. 184–85.
50. Kant's conception of inner sense will be taken up in chapter 10.
51. In a marginal comment in his copy of the first edition of the *Critique* Kant changed 'soul'

- to ‘understanding’ (LB 23: 45). Strangely enough, however, he did not incorporate this change in the second edition, where the connection between the imagination and the understanding is strongly emphasized.
52. Strawson, “Imagination and Perception,” reprinted in *Kant on Pure Reason*, edited by Ralph C. S. Walker, pp. 82–99.
  53. *Ibid.*, p. 93. It should be noted, however, that Strawson is here focusing more on Wittgenstein’s metaphor of “the echo of the thought in sight” (*Philosophical Investigations*, p. 212) than on any specifically Kantian teaching.
  54. Wilfred Sellars, “The Role of the Imagination in Kant’s Theory of Experience,” in *Categories: A Colloquium*, ed. by Henry Johnstone, p. 236.
  55. *Ibid.*, p. 243.
  56. Nevertheless, at A120–21 Kant does suggest that the role of the imagination, even in apprehension, is the generation of an image.
  57. Sellars, “The Role of the Imagination,” p. 233.
  58. Sellars, *Science and Metaphysics*, pp. 1–30.
  59. In *Science and Metaphysics*, Sellars distinguishes between an in any sense conceptual and a fully non-conceptual representing, which he argues is necessary in order to save the Kantian conception of sensibility from either incoherence or vacuity. The problem, as he sees it, is how sensibility can contribute to cognition in any but a causal manner, while remaining non-conceptual. His proposed solution is that it provides non-conceptual analogues or counterparts to our perceptual representings, which somehow guide the latter. Although deeply suggestive, I find this account of non-conceptual analogues or counterparts to conceptual representations both very obscure and of questionable Kantian credentials. It is not that Kant has no room for non-conceptual analogues to *fully* conceptual representings. We shall see, however, that these are schemata, which are already products of the imaginative synthesis and, therefore, not fully non-conceptual in Sellars’s sense.
  60. Sellars, “The Role of the Imagination,” pp. 237–40.
  61. J. Michael Young, “Kant’s View of Imagination,” *Kant-Studien* 79 (1988), pp. 140–64. See also his “Construction, Schematism, and Imagination,” *Topoi* 4 (1984), pp. 123–31.
  62. For a comprehensive statement of the hermeneutical approach to the Kantian imagination, see Makkreel, *Imagination and Interpretation in Kant*.
  63. Young, “Kant’s View of Imagination,” p. 142. Strangely enough, however, Young does not explicitly criticize Strawson on this point.
  64. *Ibid.*, pp. 153–54. Previously to this, Young suggests that even some animals have a rule-governed interpretive or discriminatory capacity of the imagination, but not the discursive capacity of the understanding. As he notes, this corresponds to the distinction that Kant draws in his *Logic* and elsewhere between *Erkenntnis* (full-fledged cognition) and mere *Kenntnis* (acquaintance) (*loc. cit.*, p. 152). This second aspect of Young’s position has been strongly endorsed by Sarah Gibbons, who emphasizes, against Strawson, the preconceptual nature of the imaginative synthesis (*Kant’s Theory of Imagination*, esp. pp. 25–26 and 30–32).
  65. On this point, see the important note on B136. Although located in the first part of the Deduction, it anticipates the central argument of the second.
  66. Keller (*Kant and the Demands of Self-Consciousness*, pp. 107–09) objected to my original

formulation of this claim on the grounds that it makes the Deduction turn on a purely ad hoc assumption that space and time must be unitary. But this is to neglect the fact that the Deduction follows the Aesthetic, where the unitary nature of space and time is supposedly established through the intuition arguments. Although one can, of course, reject the argument of the Aesthetic, one cannot reconstruct the argument of the Deduction without appealing to it.

67. This thesis has been argued recently by Waxman, *Kant's Model of the Mind*, and Longuenesse, *Kant and the Capacity to Judge*, esp. pp. 214–25. I criticize Longuenesse's reading of the note and version of this thesis in "Where Have All the Categories Gone?," pp. 74–76. Although Kant does characterize pure space and time in this manner at one point (A291 / B 347), the context makes it clear that he means by this merely that they are forms of intuiting rather than actual objects intuited, which is quite distinct from claiming that they are to be attributed to the imagination rather than to sensibility.
68. The contrary is affirmed by Longuenesse, *Kant and the Capacity to Judge*, p. 224.
69. Because I emphasize this feature of Kant's argument, my interpretation of the second part of the B-Deduction differs sharply from that of Thöle in spite of considerable agreement regarding the first part. He maintains that the argument of §26 is best read as applying not to perception as such or everything represented in space or time but only to what is determinately represented in space or time as an object of experience (See *Kant und das Problem der Gesetzmässigkeit der Natur*, esp. pp. 285–93). Thöle's reading is also affirmed by Carl, "Die B-Deduktion," in *Kant: Kritik der reinen Vernunft*, ed. by Mohr and Willaschek, pp. 206–07. As Thöle himself seems to acknowledge, however, this is more of a reconstruction of how he thinks Kant should have argued than an interpretation of the text.
70. Since here we are considering determination at the level of perception, it should be clear that in referring to the apprehension of an appearance as having a certain figure or an event as having a certain duration, we do not mean one that is specified mathematically. It is rather that an object presents itself immediately as having a certain shape and size, say, roughly, as a large oblong structure. This, then, becomes the basis for the precise mathematical determination, which involves the work of the understanding.
71. See Longuenesse, *Kant and the Capacity to Judge*, p. 199 and passim.
72. Kant refers to a "rule of apprehension" (*Regel unserer Auffassung*) in R2880 16: 557. The significance of this conception has been emphasized by Longuenesse (*Kant and the Capacity to Judge*, esp. pp. 116–18).
73. Even though Kant includes them among the dynamical categories and suggests they concern the relation of objects to the understanding (B 110), I am here leaving aside the modal categories because of their special status. Since these categories are uniquely tied to judgment, it might be argued that the first part of the B-Deduction is sufficient to establish their validity. Or, alternatively, since in the Postulates Kant states that the principles of modality are "nothing further than definitions of the concepts of possibility, actuality and necessity in their empirical use" (A219 / B 266), the claim could be that (if successful) the Deduction as a whole establishes the validity of the modal categories by showing that the others have a legitimate empirical use. Although I have a strong preference for the second option, it is not here necessary to take a stand on the issue.

74. Different versions of this reading have been offered by Prauss, *Erscheinung bei Kant*, esp. p. 277, and Keller, *Kant and the Demands of Self-Consciousness*, esp. p. 91.
75. The situation in the A-Deduction is more complex because it presents what purports to be a single line of argument in two forms (as an argument “from above” and as an argument “from below”) rather than, as in the B-Deduction, a single argument in two steps. The argument from above (A116–19), which begins with apperception, attempts to show that distinct episodes of the empirical consciousness of a single subject must be unifiable in a single universal self-consciousness. Accordingly, the argument is oriented toward the categories as conditions of experience in the sense of the connection of perceptions and is largely silent on any role they might have at the level of mere perception. Although it begins with perception and contains the previously cited note regarding the role of the imagination in perception, the latter problem is not thematized in the argument from below (A119–28) either. Nevertheless, the latter argument, the centerpiece of which is the problematic conception of transcendental affinity, turns on a slide from claims about “appearances” construed as the contents of empirical consciousness to claims about “appearances” construed in the transcendental sense as things that are known as they appear. Moreover, this parallels the slide from perception to experience in §26 of the B-Deduction. I discuss this conception of affinity and the nature of this slide in “Transcendental Affinity—Kant’s Answer to Hume?,” in *Kant’s Theory of Knowledge*, ed. by Lewis White Beck, pp. 119–27.
76. As Longuenesse observes, the Schematism and Principles not only apply the results of the Transcendental Deduction, they also help to clarify its meaning (*Kant and the Capacity to Judge*, p. 244).

## CHAPTER 8

1. Prichard, *Kant’s Theory of Knowledge*, pp. 246–47.
2. G. J. Warnock, “Concepts and Schematism,” *Analysis* 8 (1949), pp. 77–82.
3. Bennett, *Kant’s Analytic*, p. 150.
4. *Ibid.*, p. 151.
5. See, for example, Kemp Smith, *Commentary*, pp. 332–33.
6. Wittgenstein, *Philosophical Investigations*, §217. For a perceptive discussion of the relation between Kant and Wittgenstein on this point, see David Bell, “The Art of Judgment,” *Mind* 96 (1987), esp. pp. 221–24.
7. Of course, it is conceivable that one also faces the same opponent, but in that case the situation would still be significantly different, since one can assume that the opponent had learned from the experience and prepared a response.
8. The present discussion is related in interesting ways that cannot be explored here to Kant’s account of the role of a schema or rule of apprehension as containing something “universal in itself” in the formation of empirical concepts. For my treatment of the latter topic, see Allison, *Kant’s Theory of Taste*, pp. 20–30.
9. This is affirmed by Laughlan Chipman, “Kant’s Categories and Their Schematism,” *Kant-Studien* 63 (1972), p. 42, as well as by several others, including Guyer (*Kant and the Claims of Knowledge*, p. 159). Unlike Guyer, however, Chipman insists that Kant needs

- the distinction and ought not to have identified them. I agree with Chipman's claim that the distinction is needed in the case of empirical concepts but disagree with his claim that Kant does not recognize this.
10. Chipman, "Kant's Categories and Their Schematism," pp. 45–46.
  11. The expression 'discursive rule' is borrowed from Longuenesse and was already introduced in chapter 4.
  12. For a discussion of this issue see Longuenesse, *Kant and the Capacity to Judge*, p. 49. As she notes, "necessitates" here means primarily "*inescapably demands*," which is equivalent to what I have termed "normative necessity." It should also be noted that at A106 Kant himself refers to the concept of body as functioning as a "rule for intuitions," which suggests (in my terminology) that he regards it as a perceptual rule. But such confusion seems inevitable at this point, since he has not yet introduced the distinction between concept and schema.
  13. Longuenesse characterizes the rule in this sense as a "rule for sensible synthesis" and associates it with the "concept as schema" (*Kant and the Capacity to Judge*, p. 50). Although I am in substantial agreement with her on the issue, I characterize the distinction between two sorts of rules somewhat differently in order to deal with the problem internal to Kant's account of schematism, which she does not explicitly consider.
  14. Kant poses the problem in these terms in his important letter to Tieftrunk of December 11, 1797, Br 12: 224–5; 538. (See also the draft version, Br 13: 471–72.)
  15. See A137 / B176. I am here following the original version of the text (also followed by Guyer and Wood), rather than Vaihinger's emendation, which was adopted by Kemp Smith. For a discussion of this point, see Paton, *Kant's Metaphysic of Experience*, vol. 2, p. 26, n. 1.
  16. Kemp Smith, *Commentary*, pp. 335–36. Here and elsewhere in his discussion of the Schematism, Kemp Smith is closely following Ernst Robert Curtius, "Das Schematismuskapitel in der Kritik der reinen Vernunft," *Kant-Studien* 19 (1916), pp. 338–66. The line of objection sketched above is therefore intended to express the views of both Curtius and Kemp Smith.
  17. Paton, *Kant's Metaphysic of Experience*, vol. 2, p. 26, n. 1.
  18. This is noted by Curtius, "Das Schematismuskapitel," *Kant-Studien* 19 (1916), p. 348, and Kemp Smith, *Commentary*, p. 336.
  19. See Prauss, *Erscheinung bei Kant*, p. 103.
  20. Both Curtius ("Das Schematismuskapitel," pp. 348ff.) and Kemp Smith (*Commentary*, p. 336) acknowledge the suitability of the syllogistic conception for the presentation of the problematic of the Schematism but then criticize Kant for not adhering to it. In response to this criticism, I am suggesting that he does adhere to it, at least within the Schematism itself.
  21. For example, this view was upheld by Moltke Gram, who nonetheless suggests that transcendental schemata are to be understood as pure intuitions (*Kant, Ontology, and the A Priori*, esp. 128–29), and, more recently, by Gerhard Seel, "Die Einleitung in die Analytik der Grundsätze, der Schematismus und die obersten Grundsätze (A130 / B169–A158 / B197)," in *Kritik der reinen Vernunft*, ed. by Georg Mohr and Marcus Willaschek, p. 225.

22. See KpV 5: 68; 195, KU 5: 351; 225, and the draft version of Kant's December 1797 letter to Tieftrunk, Br 13: 468.
23. It should be noted here that, although an intuition is, by definition, the representation of an individual, a *determinate* intuition is always the representation of an individual under a description, that is, a particular. This applies even if, for example, what I perceive is not simply *a* house but *my* house and I perceive it as such. In that case, it is perceived under what, when reflected or conceptualized, becomes a definite description. The situation is analogous at the level of perception to the singular use of a concept at the level of judgment.
24. Kant makes a similar point at MAN 4: 478; 192.
25. For a recent expression of this criticism see Guyer, *Kant and the Claims of Knowledge*, p. 174.
26. The point is emphasized in his discussion of this issue by Alfredo Ferrarin, "Kant on the Exhibition of a Concept in Intuition," *Kant-Studien* 86 (1995), p. 163.
27. The contrary is affirmed by Greg E. Franzwa, "Space and the Schematism," *Kant-Studien* 69 (1978), pp. 149–59.
28. In the *Prolegomena*, Kant describes the concerns of both the Schematism and the closely related first part of the Phenomena-Noumena chapter as "utterly dry [*äusserst trockene*]," albeit indispensable investigations (Pro 4: 316; 108–09).
29. Kant lists only a single schema for the three categories of quantity and another one for the three categories of quality. For a plausible explanation of this procedure, see Paton, *Kant's Metaphysic*, vol. 2, pp. 63–64.
30. In the previously cited letter to Tieftrunk, Kant apparently alludes to this lacuna and suggests that it "might be remedied in another essay" (Br 12: 225; 538). Since this essay was never written, the following considerations may be viewed as a partial sketch of such a project.
31. Beck, "Can Kant's Synthetic Judgments be Made Analytic?," in *Kant: Disputed Questions*, ed. by Moltke S. Gram, pp. 241–42.
32. *Ibid.*, p. 242.
33. See Gibbons, *Kant's Theory of Imagination*, pp. 56–7, n. 12, where she criticizes my treatment of the issue in the first edition.
34. Versions of this model have been proposed by Eva Schaper, "Kant's Schematism Reconsidered," *Review of Metaphysics* 8 (1964), pp. 267–92; and, more recently, by David Bell, "The Art of Judgment," *Mind* 96 (1987).
35. This is the basic thesis of Bell, "The Art of Judgement" p. 239. In fact, according to him, "the concerns of the schematism doctrine turn out to be identical with those of the *Critique of Judgement*" (loc. cit., p. 230). In spite of the illuminating and deeply suggestive nature of much of Bell's account, I believe that this statement of an identity of concern is much too strong and seriously misleading. Apart from the vital difference between the determinative and the reflective uses of the power of judgment, which Bell barely acknowledges, Kant's concern in the third *Critique* is with the imagination in its free play, that is, its unifying form-giving activity independent of conceptual determination, whereas in the first *Critique* he is concerned with the imagination's capacity to exhibit what is thought in a specific concept and with judgment's capacity to determine the form

that such an exhibition must take. Put simply, Bell fails to distinguish sharply enough between the schematization of a concept and what Kant refers to at one point in the third *Critique* as schematizing “without a concept” (KU 5: 287; 167). For my discussion of the latter, see Allison, *Kant’s Theory of Taste*, pp. 170–71.

36. Although I am not aware of any interpretation that applies this artistic creation model to the understanding of the schematism, it is in some ways better suited to the task than the one based on the judgment of taste. Nevertheless, it, too, is inadequate for the reasons noted. For my own account of aesthetic ideas as expressions of ideas of reason in Kant’s aesthetic theory, see Allison, *Kant’s Theory of Taste*, pp. 256–62, and for their connection with genius, pp. 284–88.
37. We have already seen that the modal categories are a special case. Moreover, an adequate treatment of the connection between the categories of quantity and quality and their schemata would involve lengthy considerations that are not directly germane to our main line of argument.
38. It is sometimes claimed that only these schemata fit Kant’s definition of a transcendental schema. See, for example, Guyer, *Kant and the Claims of Knowledge*, p. 175, and Seel, “Die Einleitung,” p. 236. Without going that far, we can see that there are special difficulties in the case of the schemata of the mathematical categories. For example, Kant designates number, defined as “a representation that summarizes the successive addition of one (homogeneous) unit to another,” as the schema of quantity (A142 / B182). Although it is certainly plausible to claim that we necessarily use number to determine the quantity of things and, therefore, apply the concept to the world, it is less clear that number may be understood as a transcendental time-determination.
39. By contrast, the remaining schemata are concerned with the time-series (quantity), the content of time (quality) and the “sum total of time [ *Zeitbegriff* ] in regard to all possible objects” (A145 / B184–85). Apart from the obscurity of these labels, particularly the last, the problem with these schemata, as suggested in the preceding note, is understanding how they count as transcendental time-determinations.
40. ‘*Beharrlichkeit*’ may be translated either as ‘permanence’ or as ‘persistence’. Kemp Smith usually translates in the former manner, and Guyer and Wood in the latter. Since the Guyer and Wood translation is being used here, I shall generally follow their rendition, though on occasion ‘permanence’ will be used instead.
41. Here I am following Kemp Smith rather than Guyer and Wood, who usually translate ‘*Zugleichsein*’ as ‘simultaneity’. In my judgment, the latter gives the misleading impression that Kant is concerned only with the relation of substances at a given moment rather than throughout a stretch of time. The issue will be discussed in the subsequent chapter.
42. Kant states that the Postulates, which deal with the modal categories, are “nothing further than definitions of the concepts of possibility, actuality, and necessity in their empirical use” (A219 / B266), and he denies that they are “objective-synthetic” (A233 / B286).
43. This also applies to Kant’s remark that in the Analogies appearances are subsumed “not under the categories per se, but only under their schemata” (A181 / B224). For a discussion of this issue, see Paton, *Kant’s Metaphysic*, vol. 2, pp. 180–81.
44. For a discussion of this issue with specific reference to the Kant criticism of C. I. Lewis, see Beck, “Can Kant’s Synthetic Judgments be Made Analytic?,” pp. 235–38; the same is-

sue is treated more fully in Beck's "Lewis' Kantianism," in *Studies in the Philosophy of Kant*, pp. 108–24.

## CHAPTER 9

1. Kant refers to "modes of time" at A177 and A216 / B 262. Supposedly, each of the Analogies is correlated with one of the modes as the rule for its determination. These passages raise two exegetical issues: one concerns the meaning of the expression 'mode of time'; the other concerns the compatibility of the claim that permanence or duration, as well as succession and coexistence, are such modes with what Kant says elsewhere in the Analogies. For example, we find Kant claiming (A183 / B 226) that change (or succession) does not affect time itself, only appearances in time. He further claims in the same context that simultaneity is not a mode of "time itself" because the parts of time are not simultaneous but successive. Paton indicates the solution to this interpretive morass when he points out that these modes are not to be construed as properties or characteristics of time itself but rather as relational properties of things in time (*Kant's Metaphysic of Experience*, vol. 2, pp. 165ff.).
2. In the first edition the general principle states: "As regards their existence, all appearances stand *a priori* under rules of the determination of their relation to each other in **one** time" (A176). Although this is an accurate statement of what must be shown, the argument sketched in support of this general principle, which attempts to link it directly to the unity of apperception (as it functions in the A-Deduction), is of little relevance to Kant's actual procedure in the Analogies. We shall see, however, that Kant rectifies this in the B-version.
3. Much of the confusion and obscurity in this paragraph results from Kant's tendency to shift unannounced from the empirical to the transcendental senses of such key terms as 'appearance'. A helpful account of the terminological difficulties is provided by L. W. Beck in "A Reading of the Third Paragraph in B," in *Essays on Kant and Hume*, pp. 141–46.
4. This point is noted by numerous commentators, including Paton, *Kant's Metaphysic of Experience*, vol. 2, p. 231; Melnick, *Kant's Analogies of Experience*, p. 85; Beck, "A Reading," p. 144; and James Van Cleve, "Four Recent Interpretations of Kant's Second Analogy," *Kant-Studien* 64 (1973), pp. 69–87, esp. p. 75.
5. We shall see in chapter 12 that when in the judgment of inner sense this so-called subjective order is made into an object, it is experienced as part of the objective temporal order of the phenomenal world. Strictly speaking, there is only one temporal order because there is only one time.
6. The Newtonian position is obviously the test case here. My interpretation is based on H. G. Alexander's analysis of Newton's characterization of absolute or real time in the famous scholium to definition 8 of the *Principia* as time as it is measured by the period of revolution of Jupiter's moons and by pendulums. *The Leibniz-Clarke Correspondence*, ed. and trans. by H. G. Alexander, pp. xxxv–xxxvi. The point, then, is that not even Newtonian absolute or real time is held to be itself perceived.
7. Kant makes the point in a less ambiguous way at A197 / B 242–43.



8. The contrary thesis is the basis of Melnick's interpretation of the First Analogy; *Kant's Analogies of Experience*, pp. 58–71.
9. This line of argument is developed by W. H. Walsh in "Kant on the Perception of Time," *The First Critique*, ed. by T. Penelhum and J. MacIntosh, pp. 70–88, and in *Kant's Criticism of Metaphysics*, pp. 129–35.
10. This citation is taken from Kemp Smith, *A Commentary to Kant's "Critique of Pure Reason"*, p. 359n.
11. This is suggested by Kemp Smith, *Commentary*, p. 359.
12. Paton, *Kant's Metaphysic*, vol. 2, p. 191.
13. James Van Cleve, "Substance, Matter, and Kant's First Analogy," *Kant-Studien* 70 (1979), p. 153.
14. Although somewhat awkward and not without a certain ambiguity, I take this to be more informative than mere "change" and certainly less misleading than Bennett's recommended "existence-change" (*Kant's Analytic*, pp. 187–88).
15. What follows is suggested by the treatment of this topic by Dryer, *Kant's Solution*, pp. 353–59.
16. This argument should be contrasted with the one attributed to Kant by Van Cleve in "Substance, Matter, and Kant's First Analogy," pp. 155–57. The argument, as he constructs it, involves a non sequitur in the move from the premise that some antecedent object must have existed at  $t_1$  (if one is to experience a replacement change at  $t_2$ ), to the conclusion that the new state of affairs that comes into existence at  $t_2$  must be merely a determination or state of the object existing at  $t_1$ . The argument sketched here, however, links the state of affairs that comes into existence at  $t_2$  with a prior and contrary state of affairs (symbolized by 'non- $x$ ' and ' $x$ '), not with an antecedently existing object. It contends that if the coming into existence of the later state of affairs (the replacement change) is to be experienced, it must be contrasted with the earlier state of affairs (otherwise there would be no change), and this requires that both states of affairs (non- $y$  and  $x$ ) be experienced as successively existing states or determinations of an enduring object ( $y$ ). Thus, while I agree with Van Cleve that the argument he provides does involve a non sequitur, I see no reason to accept his reconstruction of Kant's argument.
17. Among those who maintain this in one form or another are Bennett, *Kant's Analytic*, p. 199; Dryer, *Kant's Solution*, pp. 367–68; and Melnick, *Kant's Analogies*, p. 67; Strawson, *Bounds of Sense*, pp. 128–30.
18. See, for example, A185 / B228 and A187 / B230–31.
19. This argument has been developed by Walsh; see note 9, above.
20. Critics who attack Kant on this point include C. D. Broad, "Kant's First and Second Analogies of Experience," *Proceedings of the Aristotelian Society* 25 (1926), pp. 189–210; Bennett, *Kant's Analytic*, p. 200; and Wolff, *Kant's Theory*, p. 251. The whole matter is discussed by Van Cleve, "Substance, Matter, and Kant's First Analogy," pp. 158–61.
21. A somewhat similar argument, which turns on an appeal to the Axioms of Intuition, is suggested by Carl Friedrich Weizsäcker, "Kant's 'First Analogy of Experience' and Conservation Principle of Physics," *Synthese* 23 (1971), p. 84. Support for this interpretation is also provided by R81 23: 30–31, where Kant remarks: "If substance persists while the accidents change [*wechseln*] but, at the same time, substance apart from all of its acci-

- dents is the empty substantial [*das leer substantiale ist*], what is it that persists? The only thing in experience that can be distinguished from the changing determinations is quantity [*Quantität*], and this can only be measured through the magnitude [*Grösse*] of the merely relative effect with respect to equivalent outer relations. *It therefore applies only to bodies.*"
22. For a discussion of the empirical nature of Kant's concept of matter see Walker, "The Status of Kant's Theory of Matter," in *Kant's Theory of Knowledge*, ed. by Lewis White Beck, pp. 151–56.
  23. It should be noted that this analysis enables Kant to speak of a plurality of substances, in spite of the identification of substance with matter. For a discussion of this point see Paton, *Kant's Metaphysic of Experience*, vol. 2, pp. 211–12.
  24. Van Cleve has claimed ("Substance, Matter, and Kant's First Analogy," pp. 160–61) that this argument works only for the conservation of the quantity of matter as Kant defines it. And since Kant fails to show that the quantity of matter, so construed, can be equated with mass, he fails to establish the principle of the conservation of mass. But even if this is correct, it does not affect the point that no such argument is presented in the *Critique*.
  25. Hume, *Treatise*, p. 56.
  26. Beck, "A Prussian Hume and a Scottish Kant," in *Essays on Kant and Hume*, pp. 111–29.
  27. This is also noted and emphasized by Guyer, *Kant and the Claims of Knowledge*, esp. pp. 240–42.
  28. This is affirmed in several places by Michael Friedman, who understands the function of causality in this way. See, for example, "Kant on Causal Laws and the Foundations of Natural Science," in *The Cambridge Companion to Kant*, pp. 161–97. I discuss Friedman's interpretation of the Second Analogy in *Idealism and Freedom*, chapter 6.
  29. Bird, *Kant's Theory of Knowledge*, pp. 158–59.
  30. See Paton, *Kant's Metaphysic*, vol. 2, pp. 224–25. He follows the tradition of Adickes and Kemp Smith.
  31. See Kemp Smith, *A Commentary*, pp. 375–76, who is closely followed by Wolff, *Kant's Theory of Mental Activity*, p. 373; for a more developed critique, see Ewing, *Kant's Treatment of Causality*, p. 375. A helpful restatement of these criticisms is provided by W. A. Suchting, "Kant's Second Analogy of Experience," *Kant-Studien* 58 (1967), pp. 355–69. An interesting attempt at a defense of Kant against these objections is offered by Paton, *Kant's Metaphysic*, vol. 2, pp. 254–56.
  32. For a similar assessment of the situation, see Guyer, *Kant and the Claims of Knowledge*, p. 241.
  33. See Bird, *Kant's Theory of Knowledge*, p. 155, and Melnick, *Kant's Analogies*, pp. 79–80.
  34. Typical of this are the interpretations of Wolff, *Kant's Theory*, p. 267; and Melnick, *Kant's Analogies*, pp. 80–82.
  35. The same assumption is criticized from a different point of view by Melnick, *Kant's Analogies*, pp. 81–83.
  36. In spite of the criticism cited above, this is recognized by Wolff, who writes: "The real point of the argument, as Kant makes clear later in the Analogy, is not that we must perceive *B* after *A*, but that we must represent or think *B* after *A*. Objectivity is a characteristic of cognition not of apprehension" (*Kant's Theory*, p. 268).

37. See Van Cleve, "Four Recent Interpretations," pp. 75–76, for his discussion of Ewing's interpretation; and Melnick, *Kant's Analogies*, pp. 81–82, for his critique of Paton on this point.
38. See also A111.
39. This line of objection was already formulated in Kant's lifetime by Solomon Maimon, who conceded that Kant established the *quid juris* but denied that he did so for the *quid facti*. For discussion of Maimon's critique of Kant, see Beiser, *The Fate of Reason*, pp. 285–323, and Thielke, "Discursivity and Its Discontents," esp. chapters 4 and 5.
40. The classic statement of the logic of this transcendental argument through which Kant answers Hume is provided by Beck, *Essays on Kant and Hume*, esp. pp. 131–35.
41. See, for example, Prichard, *Kant's Theory of Knowledge*, pp. 294–96.
42. Guyer, *Kant and the Claims of Knowledge*, pp. 259–62.
43. *Ibid.*, 240.
44. Reid, *Essays on the Active Powers of the Human Mind*, *Philosophical Works*, vol. 2, p. 627. For a discussion of this see Guyer, *Kant and the Claims of Knowledge*, pp. 259–60.
45. Schopenhauer, *Ueber die vierfache Wurzel des Satzes vom zureichenden Grunde*, vol. 1, *Sämtliche Werke*, ed. by J. Frauenstadt, §23, pp. 85–92.
46. Strawson, *The Bounds of Sense*, p. 137. Lovejoy, "On Kant's Reply to Hume," in *Kant: Disputed Questions*, ed. by Moltke S. Gram, pp. 284–308.
47. *Ibid.*, p. 138.
48. *Ibid.*
49. Beck, "A Non-Sequitur of Numbing Grossness," in *Essays on Kant and Hume*, pp. 151–52.
50. Guyer, *Kant and the Claims of Knowledge*, p. 252.
51. Thus, in apparent support of his reading, Guyer twice refers to A201–02 / B 246–47, where Kant refers to the objective validity of empirical judgments regarding an event (*Kant and the Claims of Knowledge*, pp. 239, 259). It should be noted, however, that on both occasions he uses it to discuss Schopenhauer's objection rather than to confirm his own underlying assumption.
52. *Ibid.*, pp. 258–59, 315, 326.
53. This formulation is taken from Brittan, *Kant's Theory of Science*, p. 189. In fairness, however, it should be noted that this does not reflect Brittan's interpretation of the argument.
54. Although he does not have in mind the precise problem with which we are at present concerned, a characterization of such a situation is suggested by Melnick with his analysis of a "closed system" (*Kant's Analogies*, p. 117). According to Melnick, "A closed system is one whose description at a certain time is governable by laws whose conditions of application constitute a description of the system at some other time." The movement of the planets is an example of such a system because "the position of the planets at any time is a function of their mass and position at other times." This can perhaps be taken as the description of a situation in which the earlier state of a "substance" is the cause or condition of its later state.
55. This is Friedman's view. See note 28, above.
56. I discuss this issue in detail in the first chapter of *Kant's Theory of Taste*.
57. See also FI 20: 209; 13.
58. See chapter 8, note 41.

59. For Kant's analysis of simultaneity, see ID 2: 401; 394.
60. Paton, *Kant's Metaphysic of Experience*, vol. 2, p. 298. More recently the same point has been made by Thöle, "Die Analogien der Erfahrung," in *Kritik der reinen Vernunft*, pp. 291–92.
61. See Baumgarten, *Metaphysica*, §211 (7, 71). The reference is noted by Paton, *Kant's Metaphysic of Experience* vol. 2, p. 143, n. 7.
62. Eric Watkins, "Kant's Third Analogy of Experience," pp. 438–40. Watkins further suggests that the same model of causality also applies to the Second Analogy, but I do not follow him here.
63. *Ibid.*, pp. 435–36.
64. A similar conclusion is reached by Paton, *Kant's Metaphysic of Experience*, vol. 2, pp. 321–22.
65. This point has been emphasized by many commentators, most recently by Margaret Morrison, "Community and Coexistence: Kant's Third Analogy of Experience," *Kant-Studien* 89 (1998), esp. pp. 266–67.
66. See, for example, B 112, A214 / B 261.
67. See Paton, *Kant's Metaphysic of Experience*, vol. 2, p. 297.
68. The counterfactual nature of Kant's claim at this point is noted by Watkins, "Kant's Third Analogy of Experience," p. 426.
69. See Paton, *Kant's Metaphysic of Experience*, vol. 2, 307. Paton does not, however, connect this with the disjunctive function of judgment.
70. This line of criticism is found in many forms, ranging from the dismissive treatments of Prichard, who thinks that Kant is simply confused on the matter (*Kant's Theory of Knowledge* pp. 304–07), and Strawson, who effectively accuses Kant of a non sequitur comparable to the one he allegedly committed in the Second Analogy (*The Bounds of Sense*, pp. 139–40), to the more sympathetic formulation of Ewing, who appeals to the distinction from the Second Analogy (*Kant's Treatment of Causality*, pp. 116–17).
71. See Schopenhauer, *The World as Will and Representation*, vol. 2, pp. 459–62. For concise useful critical discussions of Schopenhauer's views on the Third Analogy see Kemp Smith, *Commentary*, pp. 387–89, and Ewing, *Kant's Treatment of Causality*, pp. 118–19.
72. Melnick, *Kant's Analogies of Experience*, p. 96.
73. *Ibid.*, p. 103.
74. *Ibid.*, p. 104.
75. *Ibid.*, p. 105.
76. A similar criticism is raised by Watkins, "Kant's Third Analogy of Experience," p. 429.
77. Guyer, *Kant and the Claims of Knowledge*, p. 270, p. 452, n. 12. Guyer apparently has in mind Kemp Smith's example of a series of musical notes played in the reverse order (*A Commentary*, p. 386).
78. *Ibid.*, pp. 271–72.
79. *Ibid.*, p. 272.
80. This is noted by Watkins, "Kant's Third Analogy of Experience," p. 431.
81. *Ibid.*, p. 432.
82. Guyer, *Kant and the Claims of Knowledge*, p. 273.
83. Watkins, "Kant's Third Analogy of Experience," pp. 424–25.

84. Ibid., pp. 425, 436–37.
85. Ibid., p. 436.
86. The point is emphasized by Margaret Morrison, “Community and Coexistence: Kant’s Third Analogy of Experience,” *Kant-Studien* 89 (1998), pp. 257–77.

## CHAPTER 10

1. Paton, *The Categorical Imperative*, p. 233.
2. Although it is not possible to determine whether Kant had something like this in mind, his analysis suggests a distinction between things or persons and their histories, which is sometimes drawn by philosophers working in the areas of tense logic and grammar. The basic point is that individual objects, including both things and persons, do not themselves have temporal parts, extension, or even location. Temporal properties are all predicated of the histories, processes, and events connected with things and persons, but not *directly* of the things and persons themselves. See, for example, Dennis C. Holt, “Timelessness and the Metaphysics of Temporal Existence,” *American Philosophical Quarterly* 18 (1981), pp. 149–56. Holt, however, does not deal with Kant.
3. By ‘soul’ Kant means a separate immaterial substance with the capacity to think, feel, and so on. Kant suggests that anthropology abstracts from the question of whether human beings have a soul in this sense, whereas in psychology (by which Kant presumably understands the Wolffian empirical psychology of Baumgarten), such an entity is assumed. See Anthro 7: 161; 39.
4. See also KU 5: 206; 92, and Anthro 7: 153; 32.
5. Hume, *Treatise*, p. 165.
6. This point is noted by Wilfred Sellars, “. . . This I or he or it (the thing) which thinks . . .,” *Proceedings and Addresses of the American Philosophical Association* 44 (1971), p. 8.
7. See also R6354, 18: 680.
8. In R5654 18: 312, Kant refers to the object of inner sense as the “transcendental object of apperception.”
9. Elsewhere, in discussing the phenomenal nature of self-knowledge in the *Reflexionen*, Kant refers to the thesis that the contents of inner sense (apart from feelings) are derived from outer sense, but he does not suggest that this constitutes a separate argument. See, for example, R5653 18: 306–07; R5655 18: 313–16; R6313 18: 6319; 633; and LR 9–11. Thus, it may not be correct to assume that Kant took it as an independent argument in the *Critique*, though, for analytical purposes, it is useful to consider it as such.
10. For a criticism along these lines, see Hossenfelder, *Kants Konstitutions-theorie und die Transzendente Deduktion*, pp. 31, 61–63.
11. Paton, *Kant’s Metaphysic*, vol. 2. pp. 238–40.
12. See Washburn, “The Problem of Self-Knowledge and the Evolution of the Critical Epistemology, 1781 and 1787,” esp. pp. 194–215. Much of the following analysis was suggested to me by Washburn’s discussion. For a somewhat different account of self-affection, see Guenter Zoeller, “Making Sense Out of Inner Sense: The Kantian Doctrine as Illuminated by the Leningrad *Reflexion*,” *International Philosophical Quarterly* 115 (1989), pp. 263–70.

13. The significance of such a reconceptualization and change of epistemic focus is discussed at length by Prauss, *Erscheinung bei Kant*, pp. 254–321.
14. Kant's first published discussion of the problem is in NE I: 411–12; 39. The development of Kant's anti-idealism argument in the pre-critical period is discussed by Heidemann, *Kant und das Problem des metaphysischen Idealismus*, and Luigi Caranti, "The Development of Kant's Refutation of Idealism," Ph.D. dissertation, Boston University, 2001, (unpublished).
15. Although Kant reformulated the Fourth Paralogism in the second edition and has it deal with the mind-body problem rather than with external world skepticism (B 409–410), it should be kept in mind that he retained a mini-version of the argument in the section of "The Antinomy of Pure Reason" where he explicates transcendental idealism (B 519–520). In both cases, Kant's concern is to underscore the difference between his transcendental or formal idealism and the illicit empirical or material idealism. The same applies to his treatment of the problem in the *Prolegomena*, which is modeled closely on that of the first edition (Pro 4: 336–37; 91–92).
16. Guyer, "Kant's Intentions in the Refutation of Idealism," *Philosophical Review* 92 (1983), pp. 329–83, and *Kant and the Claims of Knowledge*, pp. 279–329.
17. My treatment of this issue has been influenced by the discussion of Caranti, *The Development of Kant's Refutation of Idealism*, pp. 187–93.
18. The claim that this is the only possible proof is reiterated in the *Reflexionen* (R631I–6316, 18: 607–23). The fact that there Kant sketches various proofs indicates that he viewed them as alternative expressions of the same line of argument, which is itself viewed as the only way to defeat the skeptic. Indeed, they all take the form of attempting to show that the skeptic must accept the reality of outer experience as a condition of the possibility of a minimal kind of self-knowledge that is itself presumed to be indubitable.
19. See B 70–71.
20. As Descartes made clear, the same line of reasoning also leads to a skepticism concerning other minds. Accordingly, "problematic idealism" could equally well be characterized as "problematic solipsism." But Kant completely ignores the latter dimension of the problem, focusing only on a skepticism regarding the physical world.
21. The major treatments of this topic are Lachière-Rey, *L'idéalisme kantien*, pp. 60–148, and Rousset, *La doctrine kantienne de l'objectivité*, pp. 139–61.
22. See also R5653, 5655, 18: 309, 314.
23. Rousset, *La doctrine kantienne*, pp. 148–51; Myron Gochnauer, "Kant's Refutation of Idealism," *Journal of the History of Philosophy* 12 (1974), pp. 195–206, esp. p. 198; and Richard E. Aquila, "Personal Identity and Kant's Refutation of Idealism," *Kant-Studien* 70 (1979), pp. 259–78, esp. pp. 260–61.
24. Bennett, *Kant's Analytic*, p. 205.
25. Ibid.
26. Ibid., pp. 200–14.
27. Descartes himself suggests such a temporalized understanding of the contents of consciousness when, in the Third Meditation, prior to the demonstration of the existence of God and thus still under the assumption of the demon hypothesis, he remarks in passing

that “when I think I exist now and remember in addition having existed formerly, or when I conceive various thoughts of which I recognize the number, I acquire the ideas of duration and number, which I afterward am able to apply to any other things I wish” (*The Philosophical Writings of Descartes*, ed. by J. Cottingham, R. Stoothoff, and D. Murdoch, vol. 2, pp. 30–31). Although the point that Descartes is trying to make here is not a skeptical one but rather that the mind has the resources to generate these ideas from itself (which is not the case with the idea of God), the passage does suggest that he attributed a certain diachronic dimension to the contents of the *cogito*.

28. Bennett, *Kant's Analytic*, p. 202.
29. Guyer, “Kant’s Intentions,” pp. 332–24, and *Kant and the Claims of Knowledge*, pp. 283–84.
30. Kant makes this explicit at B 278.
31. See R5653, 6315 18: 308, 618–20. The issue is discussed by Rousset, *La doctrine kantienne*, p. 155.
32. Such an objection was raised by Schulze, “*Aenesidemus oder über die Fundamente der von dem Herrn Prof. Reinhold in Jena gelieferten Elementar-Philosophie*,” reprinted in *Aetas Kantiana*, pp. 105–08.
33. This is the main question posed by Caranti, *The Development of Kant’s Refutation of Idealism*, chapter 4.
34. See R5461 18: 189.
35. Apart from Guyer’s, the most notable systematic and scholarly interpretation of the Refutation along realistic lines is by Rousset, *La doctrine kantienne*, esp. pp. 139–61. Like Guyer, Rousset emphasizes the late *Reflexionen*.
36. Guyer, “Kant’s Intentions,” pp. 330–31; *Kant and the Claims of Knowledge*, p. 281 and *passim*.
37. Guyer, “Kant’s Intentions,” pp. 360–61; *Kant and the Claims of Knowledge*, esp. pp. 310–17.
38. Guyer, “The Postulates of Empirical Thinking in General and the Refutation of Idealism,” pp. 322–23.
39. A similar point has been argued, albeit on somewhat different grounds, by Caranti, *The Development of Kant’s Refutation of Idealism*, esp. chapter 4.
40. It is noteworthy that Guyer cites the above passage on the very last page of his book (*Kant and the Claims of Knowledge*, p. 415) and suggests that the form of transcendental idealism it requires is compatible with a realistic reading of the intended conclusion of the refutation of idealism. Of course, he immediately denies the significance of this on the grounds that he has shown the irrelevance of this idealism to Kant’s transcendental theory of experience; but it is just this that I have attempted to deny.

## CHAPTER 11

1. Kant does, however, provide a hint of the role assigned to reason at A132 / B171. The initial view reflects his earlier position, as articulated in the *Inaugural Dissertation*, according to which the Latin ‘*intellectus*’ included both reason and understanding. But clear ev-

- idence of the importance that this distinction gradually came to have for Kant is provided by the *Reflexionen* stemming from the “silent decade.” The earliest of these to draw a clear distinction between the two is R4675 17: 648–53, which is dated 1775.
2. These were traditionally viewed as distinct branches of “special metaphysics,” because each dealt with a distinct object of metaphysical speculation. As such they are distinguished from “general metaphysics” or ontology, which deals with being as such or “things in general. This framework was used by Wolff and Baumgarten in the presentation of their metaphysics and underlies Kant’s critique. As we shall see, this is reflected in a division of labor between the critical functions of the Analytic and the Dialectic: the former providing a critique of general, and the latter of special metaphysics.
  3. This marks a reversal of the order of topics in Kant’s own Introduction to the Dialectic. For Kant begins with a discussion of illusion and moves from this to his preliminary account of reason. But since this order seems to have been dictated by Kant’s characterization of dialectic as the “logic of illusion” rather than by any systematic considerations, it seems better to begin with his account of reason, which underlies the doctrine of illusion.
  4. Thus, Bennett suggests that Kant should have denied that there are any principles in this sense. (See *Kant’s Dialectic*, p. 262.)
  5. Kant sometimes characterizes the former as a “collective” and the latter as a “distributive unity.” (See A 582 / B 610 and A 644 / B 672). Moreover, this distinction is foreshadowed in the *Nachlass*. For example, in R4757 17: 703, Kant relates distributive unity to principles of the possibility of experience; and in R4760 17: 711, he connects collective unity with the unity of the whole use of reason. Like R4675 (see note 1), these stem from the period of Kant’s earliest attempts to draw a sharp distinction between the understanding and reason.
  6. Kant assigns immediate inferences to the understanding. Basically, these are the inferences that can be drawn on the basis of the Aristotelian square of opposition. See JL 9: 115–119; 610–14.
  7. Since ‘*Vernunftschluss*’, which literally means ‘inference of reason’, is the term that Kant uses to refer to a syllogism, they two will usually be treated here as interchangeable.
  8. The situation is complicated, however, by the fact that in his *Logic* Kant notes that only the categorical syllogism is really an inference of reason, since only it has a middle term (JL 9: 129; 623). Thus, he also suggests that all hypothetical syllogisms are really immediate inferences based on an antecedent proposition and a consequent proposition (JL 9: 129; 623), whereas all disjunctive syllogisms involving more than two disjuncts are polysyllogistic (JL 9: 130; 624). Strangely enough, Kant ignores this in his derivation of the ideas.
  9. It is important to keep in mind that the class of ideas of reason for Kant is broader than the class of transcendental ideas, since it includes such practical ideas as virtue and a republican constitution, which function as archetypes. Kant underscores this point by prefacing his account of transcendental ideas with a discussion of ideas in general (A 312–20 / B 368–77). It is, however, only the former with which we are here concerned.
  10. The latter is a central theme in the work of Longuenesse, *Kant and the Capacity to Judge*.
  11. In the *Prolegomena* Kant seems to entertain the possibility that the ideas are innate, but he suggests that if they are not their source must lie in the same act of reason that is ex-



pressed in the various forms of syllogistic inference (Pro 4: 430; 84). The characterization of them as inferred concepts in the *Critique*, however, indicates that he did not view them as innate.

12. Walsh, *Kant's Criticism of Metaphysics*, p. 173.
13. Ibid., pp. 174–75.
14. For a particularly harsh treatment of the idea of the soul and the categorical syllogism from this point of view see Bennett, *Kant's Dialectic*, p. 281.
15. The parallel is nicely expressed by Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 133–37.
16. At B111 Kant characterizes the category of allness or totality as “nothing other than plurality considered as a unity.”
17. In this context, it is noteworthy that Kant not only distinguished sharply between a merely comparative universality obtained through induction and true universality (B3–4), but also connected the former with the unity of the understanding and the latter with the unity of reason. See, for example, R4759 17: 709, where he refers to the universal of understanding as merely “*totalitas secundum quid*,” in contrast to absolute totality; and R5553 18: 224.
18. In a *Reflexion* (dated between 1778 and the early 1780s) Kant writes: “The concept of the totality of the synthesis according to the categories of relation is the pure concept of reason” (R5555 18: 231).
19. See Bennett, *Kant's Dialectic*, pp. 262–64; and Walsh, *Kant's Criticism of Metaphysics*, pp. 173–74.
20. Kemp Smith, *Commentary*, p. 450.
21. See Grier, *Kant's Doctrine of Transcendental Illusion*, p. 136.
22. I have emphasized the independence of the conditions of intuition because of Kant's reference to “the synthetic unity of intuition under the categories.” As in the first part of the B-Deduction, here we may take Kant to be referring to “intuition in general,” which is precisely what is involved in the pure use of a category, as contrasted with a logical function on the one hand and a schema on the other.
23. Kant characterizes his project as a “subjective derivation of [or “introduction to”] them [these ideas] from the nature of our reason” (A336 / B393). The uncertainty in the characterization stems from the disagreement between the Academy text, which has *Anleitung* [introduction] and Erdmann's emendation, which has *Ableitung* [derivation]. Although the former is certainly possible, I believe that the latter makes more sense, since it corresponds with what Kant is actually attempting to do.
24. This is particularly clear in Kemp Smith, *Commentary*, p. 454. See also England, *Kant's Conception of God*, pp. 117–18.
25. See, for example, Bennett, *Kant's Dialectic*, p. 258.
26. Such a view is suggested by Strawson, *The Bounds of Sense*, pp. 159–60; Bennett, *Kant's Dialectic*, p. 283; Walsh, *Kant's Criticism of Metaphysics*, p. 176.
27. Evidence from the *Nachlass* suggests that the tripartite division of the Dialectic was a relatively late development, the first expressions of which are found in *Reflexionen* 5552–5555, 18: 218–235, which are dated from 1778 and 1779. Earlier *Reflexionen* (4756–4760, 17: 699–713) indicate that Kant initially intended the Dialectic to be coextensive with his

- treatment of the Antinomies. For further discussion of this see Guyer, “The Unity of Reason.” Within the *Critique*, as is often pointed out, Kant suggests that the Second Antinomy is relevant to the question of the simplicity of the soul (A463 / B491), and an argument for the existence of God is taken up in the Fourth. Finally, in the two later *Critiques* Antinomy and Dialectic are explicitly identified.
28. Kant points this out in this Remark on the Thesis of the Fourth Antinomy (A456–58 / B484–86).
  29. The following discussion is heavily indebted to Grier, *Kant’s Doctrine of Transcendental Illusion*, esp. pp. 101–30. In my judgment, her account is not only the most thorough but also the most illuminating treatment of this complex and relatively neglected topic to be found in the literature. This is not to say, however, that she would endorse all aspects of the present account.
  30. Kant had already appealed to such a model to explain metaphysical error (or illusion) in 1766 (Tr 2: 342–48; 329–35). For a discussion of this see Grier, *Kant’s Doctrine of Transcendental Illusion*, pp. 32–45.
  31. Kant there defines a subreptic axiom as “a *hybrid* axiom, which tries to pass off what is sensitive as if it necessarily belonged to a concept of the understanding” (ID 2: 412; 408). For a discussion of Kant’s account of subreption in the *Dissertation*, see Grier, *Kant’s Doctrine of Transcendental Illusion*, pp. 57–64.
  32. Bennett, *Kant’s Dialectic*, p. 267, dismisses this account of error as in conflict “with everything else Kant says on the subject, unless we re-interpret it intolerably.” The account offered here is an attempt to show that this is not the case. Once again, I am indebted to Grier, *Kant’s Doctrine of Transcendental Illusion*, pp. 102–10, though I go beyond her in attempting to incorporate the critique of Leibnizian ontology in this account.
  33. See, for example, Bennett’s dismissive treatment of this distinction (*Kant’s Dialectic*, p. 267).
  34. See Grier, *Kant’s Doctrine of Transcendental Illusion*, pp. 117–30.
  35. Walsh, *Kant’s Criticism of Metaphysics*, p. 173.
  36. This analogy is suggested by Camilla Serck-Hanssen, “Transcendental Apperception: A Study of Kant’s Theoretical Philosophy and Idealism,” Ph.D. dissertation, University of California, San Diego, 1996, p. 220 (unpublished).
  37. In the first edition of this work I failed to distinguish sharply between transcendental realism and transcendental illusion. And once again I am indebted to Grier for pointing out to me the importance of this distinction for understanding Kant’s procedure in the Dialectic (see *Kant’s Doctrine of Transcendental Illusion*, pp. 101, 143, 151).

## CHAPTER 12

1. Bennett, *Kant’s Dialectic*, p. 66, adopts the phrase “Cartesian basis” to characterize the procedure of rational psychology.
2. The major exception to this is Grier, *Kant’s Doctrine of Transcendental Illusion*, chapter 5.
3. Bennett, *Kant’s Dialectic*, p. 72.
4. Ameriks, *Kant’s Theory of Mind*, p. 68.

5. Something like this interpretation of the fallacy is suggested, though not endorsed, by Bennett, *Kant's Dialectic*, pp. 74–75.
6. See Grier, *Kant's Doctrine of Transcendental Illusion*, esp. pp. 58–59.
7. It is important to keep in mind that by an “empirical use” of a category Kant understands one involving its schema, which is not the same as its use in a merely empirical judgment.
8. See, for example, A400, where Kant remarks that the I of apperception is “substance in concept, simple in concept, etc., and thus all these psychological theorems are indisputably correct.” As he goes on to point out, the crucial question is not whether this I is “substance in concept” but whether it is actually a substance; and the latter does not follow from the former.
9. I am once again following the lead of Grier, who provides a similar analysis of the formal fallacy, albeit in considerably more detail and somewhat different terms. (See *Kant's Doctrine of Transcendental Illusion*, pp. 153–60.)
10. What Kant actually says is that one might name this illusion the subreption of hypostatized consciousness. This is even more confusing, since it suggests the identification of the illusion with the subreption, which is directly contrary to Kant's sharp distinction between illusion and error.
11. See also A395; A582–84 / B610–11; A693 / B721. The second passage will be discussed in chapter 14.
12. The limitation of this reciprocity thesis to special metaphysics is necessary because, as we have seen, ontology for Kant involves subreption; but since the latter does not require the positing of a distinct set of transcendent objects, it does not involve hypostatization. That is also why it does not involve transcendental illusion.
13. At the end of this discussion (A403–05) Kant adds a brief justification of the systematic ordering of the Paralogisms on the basis of the table of categories; but I shall ignore this here because it seems of little relevance to the actual critique of rational psychology.
14. Kant explicitly connects the unity of apperception with the unconditioned required by reason at A398. It is, however, an underlying feature of the entire discussion.
15. This is explicitly affirmed by Grier, *Kant's Doctrine of Transcendental Illusion*, p. 144.
16. I take it that here Kant is not claiming that the objective necessity is attributed to things in themselves rather than merely appearances, as if that were the illusion. The point is rather that what is merely a subjective requirement of thought is viewed as if it were objective, that is, as a determination of things.
17. The passage is cited by Grier, *Kant's Doctrine of Transcendental Illusion*, p. 146.
18. This connection between transcendental realism and transcendental illusion is clearly made by Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 150–52; and it constitutes an important corrective to my earlier view, which basically identified the two. As Grier argues at several points, this identification cannot be made, since Kant holds that the illusion is unavoidable but transcendental realism is not.
19. Yet see A380, where Kant connects the taking of appearances as things in themselves with the metaphysical positions of both materialism and spiritualism.
20. See, for example, Tr 2: 327–28; 314–15. For a discussion of the early history of Kant's views on this topic, see Ameriks, *Kant's Theory of Mind*, pp. 27–37.

21. For an analysis of the use of this argument by Leibniz (called the “unity argument”) see Margaret Wilson, “Leibniz and Materialism,” *Canadian Journal of Philosophy* 4 (1974), pp. 495–513.
22. For a discussion of this issue, see Wilson, “Leibniz and Materialism,” p. 511, and Allison, *Idealism and Freedom*, pp. 97–98.
23. This is affirmed by Kitcher, *Kant’s Transcendental Psychology*, p. 211. For a critique of Kitcher on this point, see Grier, *Kant’s Doctrine of Transcendental Illusion*, p. 167.
24. This procedure accords with the distinction Kant draws between dogmatic, critical, and skeptical objections (A388–89). As Kant there notes, both dogmatic and skeptical objections are directed (albeit in different ways) against the propositions under consideration and, therefore, presume some insight into the nature of the object. By contrast, critical objections, which are the kind Kant is advancing in his discussion of the Paralogisms, are directed only against the proposed proofs of propositions. Consequently, they are free from any such presumption.
25. See chapter 7 for a discussion of the analyticity of the principle of apperception in the B-Deduction and its relation to the treatment of apperception in the A-Deduction.
26. The final portion of the Paralogism (A356–61) is devoted to a complex and, at times, confusing demonstration of the uselessness of the unity argument for establishing the immateriality of the soul, which is clearly the desired goal. For a good discussion of some of the issues involved, including distinctions between various senses of ‘materialism’ and its opposite, see Ameriks, *Kant’s Theory of Mind*, esp. pp. 35–47, 73–76. Unlike Ameriks, however, I view this discussion as an appendix to the argument rather than as its core.
27. I discuss the connections between the argument of the Third Paralogism and Locke’s theory of personal identity in Allison, “Locke’s Theory of Personal Identity,” in *Locke on Human Understanding*, ed. by I. C. Tipton, pp. 105–22.
28. See Hume, *Treatise*, p. 171.
29. Elsewhere Kant distinguishes between moral and psychological personality. Kant defines the former as the “freedom of a rational under moral laws” (which is the condition of imputability) and the latter as “the ability to be conscious of one’s identity in the different conditions of one’s existence” (MS 6: 223; 378).
30. Shoemaker, *Self-Knowledge and Self-Identity*, p. 135.
31. Consequently, I strongly reject Bennett’s characterization of the Third Paralogism as embodying Kant’s “fumbling attempt” to say something about specific judgments of self-identity (those expressible in the form ‘It was I who was *F* at *t*’), namely, “In the Cartesian basis I cannot know that someone was *F* and wonder whether it was I” (*Kant’s Dialectic*, p. 95). In my view, there is nothing “fumbling” in Kant’s analysis, and it has nothing to do with the reliability of memory.
32. Kant also added a critique of Mendelssohn’s argument for immortality (B413–18) and some reflections on the practical significance of his analysis (esp. B423–26), which are closely related to themes taken up in the second *Critique*. Moreover, as Kant himself notes, a major concern of his revision was to respond to what he regarded as “misunderstandings” of his initial account (Bxxxviii). Although Kant does not identify the latter, it seems probable that he had in mind the critiques of his treatment of the self, particularly with respect to the phenomenal-noumenal distinction, mounted by H. A. Pistorius and

- A. H. Ulrich. For discussions of these points see Kemp Smith, *Commentary*, p. 467; and R. P. Horstmann, "Kant's Paralogisms," *Kant-Studien* 84 (1993), p. 412. I discuss Pistorius's critique insofar as it bears on the issue of freedom in *Kant's Theory of Freedom*, pp. 29–53.
33. *The Philosophical Writings of Descartes*, vol. 2, p. 54.
  34. In the discussion of the A-Paralogisms the transcendental realist's problem was described as an inability to draw the appropriate subjective-objective distinction, whereas it is here said to concern the sensibility-understanding (intuition-concept) distinction. Given the discursivity thesis, however, these come to much the same thing.
  35. Williams, "The Certainty of the Cogito," in *Descartes: A Collection of Critical Essays*, ed. by Willis Doney, p. 95. Spinoza's statement is from *Principles of Cartesian Philosophy*, vol. 1, Prolegomena.
  36. Leibniz, *New Essays on Human Understanding*, trans. by P. Remnant and J. Bennett, book 4, chap. 7, §7 (p. 411).
  37. *Ibid.*
  38. It also underlies Kant's famous remark in his critique of the ontological argument, "Being is obviously not a real predicate, i.e., a concept of something that could add to the concept of a thing. It is merely the positing of a thing or of certain determinations in themselves" (A 598 / B 626). We shall return to this issue in chapter 14.
  39. In support of this rejection it is customary to cite Descartes's claim, "When someone says 'I am thinking, therefore I am or I exist,' he does not deduce existence from thought by means of a syllogism, but recognizes it as something self-evident by a simple intuition of the mind" (Response to the Second Set of Objections, in *The Philosophical Writings of Descartes*, vol. 2, p. 100). The interpretive issue, however, is not nearly as clear cut as this and similar passages might suggest. For example, according to Anthony Kenny (*Descartes: A Study of His Philosophy*, p. 51), "The premise '*cogito*' in conjunction with the presupposition that it is impossible for that which is thinking to be non-existent yield the conclusion '*sum*.'" Kenny also points out that the only thing that is required to interpret this inference as a simple syllogism is the acceptance of the thesis that 'existence' is a predicate. Since Kant criticizes Descartes's version of the ontological argument precisely on the grounds that it treats 'existence' as a (real) predicate, it is certainly reasonable for him to construe the *cogito* inference in the same way. Moreover, given this reading Kant is perfectly correct in his contention that it implies that everything that thinks exists necessarily, "for in that case the property of thought would render all beings who possess it into necessary beings." The issue is discussed by Williams, "The Certainty of the Cogito," p. 94.

## CHAPTER 13

1. Although Kant refers to four distinct Antinomies, just as he does to Paralogisms, it is often pointed out that the term *antinomy* is used throughout in the singular and refers to the general problem that arises in the attempt to reason about the sensible world as a whole. See Heimsoeth, *Transzendente Dialektik*, p. 199; Martin, Kant's *Metaphysics and Theory of Science*, pp. 62–64; and Grier, *Kant's Doctrine of Transcendental Illusion*, p. 173.

2. This recalls Hume's "dangerous dilemma," defined as a choice "betwixt a false reason and none at all," in which he finds himself caught in the conclusion to book I of the *Treatise* (p. 174). On the possible connection of this with Kant via Hamann, see Manfred Kuehn, "Hume's Antinomies," *Hume Studies* 8 (1983), pp. 25–45.
3. See Pierre Bayle, *Historical and Critical Dictionary*, s.v. "Zeno of Elea."
4. Kant describes how empiricism becomes dogmatic at A471–72 / B499–500. See also A781 / B809.
5. This dialectical syllogism and its fallacy must be distinguished from the arguments for the theses and antitheses of the various Antinomies. The latter are not presented in syllogistic form and, though the matter is highly controversial, Kant clearly regards these proofs as valid. Thus, in the *Prolegomena*, where Kant does not discuss the specific arguments, he "vouches" for their correctness. (Pro 4: 340; 131).
6. Since a so-called hypothetical syllogism has no middle term, here the fallacy cannot be characterized as that of an ambiguous middle, but it is analogous to it.
7. This apt characterization of the subject matter of the First Antinomy is taken from Bennett, *Kant's Dialectic*, p. 114.
8. See Kemp Smith, *Commentary*, p. 479. He appeals to this apparent inconsistency in support of his "patchwork" reading, according to which the account of the cosmological ideas represents Kant's earlier position, which identified the Dialectic with the Antinomy.
9. See Schopenhauer, *The World as Will and Representation*, vol. 1, p. 498, Kemp Smith, *Commentary*, p. 495; Bennett, *Kant's Dialectic*, p. 241; Wilkerson, *Kant's Critique of Pure Reason*, p. 133.
10. For a helpful treatment of this issue, see Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 218–229. She correctly suggests that the God of Spinoza nicely fits the conception that Kant presents here. Historically, however, it seems more likely that Kant had in mind the God of Newton as interpreted by Clarke in the correspondence with Leibniz.
11. This may also be understood in terms of the contrast between the Mathematical and Dynamical Principles. Recall that whereas the former deal with appearances *qua* intuited, the latter refer to what is *thought* rather than intuited. Although this did not play any role in the Principles, it opens up the possibility that something might be thought that cannot be given in possible experience.
12. This distinction, which is essential to Kant's overall philosophical project because it allows conceptual space for freedom, has been called into question by Kemp Smith (*Commentary*, p. 511) and Walsh (*Kant's Criticism of Metaphysics*, pp. 211–12), among others. The problem concerns mainly the Second Antinomy, which is sometimes thought to accord more with the dynamical than the mathematical model. Thus, Walsh suggests that rather than arguing that both thesis and antithesis are false, since neither the claim that nothing exists in the world save the simple and what is composed of simples (the thesis), nor the claim that there exists nowhere in the world anything simple (the antithesis), can withstand the *reductio* mounted by the proponent of the opposing view, it might be argued that thesis and antithesis may both be true. For what the antithesis, on Walsh's reading, rules out is merely the possibility of finding anything simple in the course of empirical inquiry, which leaves room for locating it in the noumenal realm. Walsh's conclusion

is correct but beside the point, since it ignores the cosmological nature of the question at issue. Admittedly, the latter seems open to challenge on the grounds that, unlike the First Antinomy, the Second does not appear to be concerned with the world as a whole. But this is to misconstrue the question with which it is concerned. What makes this question cosmological, indeed mathematical, is that it concerns *all composition*, that is, it addresses the very concept of a spatial composite as such. Consequently, here, as in the First Antinomy, the appeal to something non-sensible would be out of place, since it could not enter into such a composition. For an account of the connections between the two Mathematical Antinomies, see Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 209–14.

13. It should be clear from the context that the disjunction is here taken in the exclusive sense. The antithesis position is symbolized as *I* (for infinite) rather than as  $-F$  (not finite) because it argues for an actual infinite, rather than claiming merely that the world cannot be considered as finite in relevant respects. In fact, we shall see that this is precisely what is wrong with the position.
14. This is clearly stated by Al-Azm, *Origins of Kant's Argument in the Antinomies*, p. 8. Strangely enough, however, this important point is overlooked by Guyer. See *Kant and the Claims of Knowledge*, pp. 407–08.
15. Calling attention to this fact is the major contribution of Al-Azm, *The Origin of Kant's Argument*. Unfortunately, however, he tends to neglect the idealized nature of the dispute, creating the impression that throughout the Antinomy Kant is doing little more than juxtaposing the Leibnizian and Newtonian positions (the latter as represented by Clarke). For an insightful and balanced discussion of this issue see Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 182–83 and *passim*.
16. A version of this familiar line of criticism is advanced by Guyer, *Kant and the Claims of Knowledge*, p. 407. According to Guyer this and, indeed, all the antinomial arguments “turn on purely epistemological conclusions” concerning what is impossible to represent or conceive. Consequently, he suggests that they can only yield the metaphysical results at which the arguments presumably aim by presupposing either “whatever is true of things that really exist has to be knowable,” which Kant explicitly denies, or the truth of transcendental idealism. This is to assume, however, that Kant is presenting his own arguments rather than those of the rational cosmologist, who is committed to transcendental realism and, therefore, in the grip of transcendental illusion.
17. Kemp Smith, *Commentary*, p. 485.
18. Russell, *Our Knowledge of the External World*, pp. 160–61.
19. *Ibid.*, p. 161.
20. *Ibid.*
21. Strawson, *The Bounds of Sense*, p. 176.
22. Moore, *Some Main Problems of Philosophy*, p. 179.
23. *Ibid.*; and Bennett, *Kant's Dialectic*, pp. 118–19.
24. Moore, *Some Main Problems*, p. 181.
25. Such an argument was offered by Martin Knutzen. For a discussion of this issue see Cohn, *Geschichte des Unendlichkeitsproblems in abendländischen Denken bis Kant*, p. 215.
26. We can see from this the inappropriateness of such a criticism as that of Swinburne, *Time and Space*, p. 282. According to Swinburne, the meaningfulness of the infinitistic

thesis can be defended on the grounds that we can understand what it means to say that there is no limit to the succession of events, even if we cannot count them. This is quite correct but irrelevant; for the question still arises whether we are to understand this thesis, with Kant, as expressing a regulative idea to the effect that we can always conceive of further conditions, or with the “dogmatist,” as asserting the existence of an actual infinity of conditions (past events).

27. I am indebted to William McKnight for this formulation. This concept of the infinite can be formally defined in the following way:  $S$  contains infinitely many members iff for every  $n$ ,  $S$  contains more than  $n$  members (where  $n$  is a natural number).
28. As McKnight has also noted, the definition of the infinite given in the preceding note is equivalent to the Cantorian definition according to which  $S$  has infinitely many members iff there is a proper subset of  $S$ ,  $S'$ , and a relation  $R$ , such that  $R$  sets up a one-to-one correlation between  $S$  and  $S'$ . The issue is also discussed with specific reference to Russell's formulation by Moltke S. Gram, “Kant's First Antinomy,” *Monist* 51 (1967), pp. 499–518, esp. p. 514. As before, the key is to take ‘number’ in the Kantian characterization of the mathematical infinite as equivalent to ‘natural number’. For to say, with Kant, that the infinite is greater than any (natural) number, is to say that it is reflexive and non-inductive; in short, that it cannot be arrived at by a process of counting which begins with the natural numbers.
29. The expressions *totum analyticum* and *totum syntheticum*, as well as the claim that space and time are the former and bodies the latter, occur in *Reflexion* 393, *Reflexionen Kants zur kritischen Philosophie*, ed. by Benno Erdmann, p. 121. Although Al-Azm does not refer to this *Reflexion*, an account of this distinction and its relevance to Kant's argument is found in Al-Azm, *Origins of Kant's Arguments*, pp. 9–22.
30. In another passage (*Critique of Pure Reason*, A428 / B 456n.), Kant remarks that an indeterminate quantum could be intuited as a whole (*als eine Ganzes*), if it were given or enclosed within limits. Since in that case the limits themselves determine the totality, the latter could be grasped without having to be constituted in thought. The point, however, is that the proponent of the infinitistic position cannot claim this of the world. As infinite (*ex hypothesi*) it is neither indeterminate nor limited. Consequently, in this case, a synthesis would be both necessary and impossible.
31. An interesting misinterpretation of Kant on this point is found in Swinburne, *Time and Space*, pp. 282–83. Although he recognizes that the difficulty that Kant is raising involves talking about the universe as a whole, he mistakenly assumes this to be a worry about the general impossibility of making a claim about all the members of a class. Having construed it in this way, he is obviously able to dismiss it easily, simply by pointing out that we can talk about, for example, all swans. Clearly, however, this has nothing to do with Kant's argument, for talk about the universe as Kant construes it is about a high-order individual, not about the members of a class.
32. See *Critique of Pure Reason*, A410–11 / B 437. The issue is discussed by Ewing, *A Short Commentary on Kant's "Critique of Pure Reason,"* p. 210.
33. Unfortunately, I was guilty of this confusion in the first edition of this book. It is a confusion because it (falsely) implies that the assumption of an absolute beginning is meaningless rather than simply false.



34. Actually, Leibniz argues against the thesis that God could have created the universe at an earlier (or later) moment in cosmic time. See his fourth letter (p. 38) and his fifth letter (pp. 75–77).
35. Strawson, *Bounds of Sense*, pp. 177–79.
36. Keith S. Donnellan, “Reference and Definite Descriptions,” *Philosophical Review* 75 (1968), pp. 281–304.
37. Bennett, *Kant’s Dialectic*, p. 160.
38. *Ibid.*, p. 161.
39. *Ibid.*
40. *Ibid.*
41. Although Kant does not draw precisely this distinction, he does remark in VR 28: 1095 that the creation of the world cannot be regarded as an event in time, since time first began with creation.
42. See Strawson, *Bounds of Sense*, pp. 177–78; and Grier, *Kant’s Doctrine of Transcendental Illusion*, p. 190n., who echoes Strawson on this point.
43. Kant distances himself from this position in the note to A 521 / B 549, where he distinguishes his “critical” resolution of the Antinomy from the “dogmatic” proof of the antithesis.
44. For my views on the connections between the Third Antinomy and Kant’s treatment of the free-will problem, see Allison, *Kant’s Theory of Freedom*, pp. 11–53, and “The Antinomy of Pure Reason, Section 9,” pp. 475–83. The latter expresses my latest thoughts on the matter, which have evolved over the years.
45. See Heinz Röttges, “Kants Auflösung der Freiheitsantinomie,” *Kant-Studien* 65 (1974), p. 37, and Ortwein, *Kants problematische Freiheitslehre*, p. 24.
46. Schopenhauer, *The World as Will and Representation*, vol. 1, pp. 497–98.
47. Kemp Smith, *Commentary*, p. 493.
48. Bennett, *Kant’s Dialectic*, p. 185.
49. *Ibid.*, p. 186.
50. Ewing, *Short Commentary*, p. 218 (cited by Bennett, *Kant’s Dialectic*, p. 184).
51. Bennett, *Kant’s Dialectic*, p. 186.
52. See Leibniz’s Fifth Letter to Clarke, esp. secs. 1–20, *The Leibniz-Clarke Correspondence*, pp. 55–60. Much to the dismay of Clarke, Leibniz there extends this principle to the divine will itself, claiming that it is determined (though not necessitated) by what the divine intellect recognizes as the best. Moreover, in insisting on the strict (unlimited) universality of this principle, he explicitly rules out the spontaneity affirmed by Clarke to be necessary to the understanding of divine (or human) agency.
53. *Ibid.*, p. 95.
54. This is maintained by Al-Azm, *The Origins of Kant’s Argument*, pp. 92–93.
55. A somewhat similar analysis is offered by Walsh, *Kant’s Criticism of Metaphysics*, pp. 204–05. Walsh, however, does not focus specifically on the Third Antinomy, and he sees the contradiction in a tension within a principle of sufficient reason to which both Leibniz and Newton / Clarke pay lip service, rather than within Leibniz’s own understanding of the principle. More important, in light of his wholesale dismissal of Kant’s theory of reason, Walsh overlooks the connection between this principle and  $P_2$ . We shall see that

this underlies and vitiates his critique of Kant's indirect argument for transcendental idealism.

56. For contemporary formulations of this line of criticism, which can be traced back to Hegel, see Röttges, "Kant's Auflösung", p. 36, and Ortwein, *Kants problematische Freiheitslehre*, pp. 24–26.
57. Here I am following the rendering of Kemp Smith rather than of Guyer and Wood, who render "*dem vorhergehenden eben derselben Ursache*" as "the cause of the previous one," since it seems to make better sense of the argument. What is at issue is the connection between a dynamical beginning and the previous state of the agent, not the cause of that state.
58. Schopenhauer, *The World as Will and Representation*, vol. 1, p. 498.
59. Strawson, *The Bounds of Sense*, pp. 208–09.
60. See Röttges, "Kants Auflösung," pp. 45–48, and Ortwein, *Kants problematische Freiheitslehre*, pp. 26–27.
61. See especially Gr 4: 446–47; 94–95.
62. See A506 / B534.
63. The former of these phrases in quotation marks is the Kemp Smith and the latter the Guyer and Wood rendering of "*aufgegeben*." I shall here follow the former for stylistic reasons.
64. Kant here refers to it as "a logical postulate of reason" (A498 / B526).
65. See note 3.
66. Correlatively, if the thesis were merely to deny that the world is infinite, the two claims would be not only compatible but also from Kant's point of view, correct.
67. The thesis says more than it knows when it moves from the denial of the infinity of the world to the assertion of its finitude.
68. Here it should be kept in mind that all that is required for the idealism argument are the negative results of these arguments, that is, their demonstrations that the world cannot be infinite and that it cannot be finite. Their further positive conclusions, which are based on the assumption that the opposed views are genuine contradictories, are clearly illicit from Kant's standpoint. But this has no bearing on the soundness of the *modus tollens*.
69. This line of criticism is sketched by Gram, "Kant's First Antinomy," pp. 509–12.
70. One can, of course, also claim that the average family has, say, 1.7 children; but this is clearly a statistical thesis, which does not refer to any *particular* family.
71. In fact, on occasion, Kant argues directly from the self-contradictory nature of this idea of the world to a resolution of the antinomial conflict. See Pro 4: 342; 132–33, Fort 20; 321; 403. I discuss this line of argument in more detail in "Kant's Refutation of Realism," *Dialectica* 2/3 (1976), pp. 224–53.
72. Walsh, *Kant's Criticism of Metaphysics*, p. 205.
73. *Ibid.*, p. 204.
74. *Ibid.*, p. 202. See also p. 205, where Walsh suggests that Kant's attempt to show that transcendental realism is false, argues that it must at least be intelligible.
75. *Ibid.*, pp. 205–06.
76. *Ibid.*, p. 205.

77. See A 509 / B 537 for a clear statement of this point.
78. This should not, however, be taken as claiming that the entire function of transcendental idealism is therapeutic. Clearly, Kant also sees it as the essential means (on my reading, through the conception of epistemic conditions) of accounting for the possibility of synthetic *a priori* knowledge, which is the point emphasized in the Aesthetic and Analytic. Nevertheless, the therapeutic function is likewise important and predominates in the Dialectic.
79. Guyer, *Kant and the Claims of Knowledge*, p. 386.
80. Ibid., pp. 385–415. In a footnote on p. 462, Guyer criticizes my account of the indirect proof in the first edition of this book. I must confess, however, that I am unable to recognize my analysis in the characterization he provides of it.
81. For a critique of Guyer's analysis of the development of Kant's views on the nature and resolution of the antinomial conflict, see Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 191–94.
82. Perhaps Kant's clearest statement of the connection between the Antinomy and the theocentric perspective is in his letter to Herz discussing the work of Solomon Maimon. See B II: 70; 316.
83. It is because they tend to view this proof as an intended argument for something like the latter that critics do not take it seriously. See, for example, Strawson, *The Bounds of Sense*, pp. 199–206.
84. I am here altering the Guyer and Wood translation, which renders this as “deceptive illusion.” Although this certainly sounds better in English, it erroneously suggests that Kant is stating that transcendental illusion arises from the misinterpretation of our empirical concepts. As we have seen, the relationship is rather the reverse.

## CHAPTER 14

1. As is often the case, the locus classicus of this line of criticism is Schopenhauer, *The World as Will and Representation*, vol. 1, pp. 508–09. And once again, Schopenhauer was followed closely by Kemp Smith, *Commentary*, pp. 522–25. A similar critique is made by England, *Kant's Concept of God*, pp. 120–25.
2. This is affirmed by Strawson, *The Bounds of Sense*, pp. 221–23.
3. Kant states that it “rests on the principle of contradiction” (A 571 / B 579), but as Allen Wood points out, it is equivalent to it (*Kant's Rational Theology*, p. 42).
4. This point is most forcefully expressed by Peter Rohs, “Kant's Prinzip der durchgängigen Bestimmung aller Seienden,” *Kant-Studien*, 69 (1978), esp. pp. 170–72.
5. The sense of ‘possibility’ involved here is an important issue. It is clear that Kant is referring to real rather than merely logical possibility at this point, but it is not clear how this real possibility is to be understood. Following Wood (*Kant's Rational Theology*, pp. 44–50), I take Kant to be referring here not to the sense of real possibility as agreement with the formal conditions of experience presented in the Postulates (A 218 / B 265), but rather to what he there understands by “absolute possibility,” namely, that which is possible in all respects, as opposed to a contingent experience. Significantly, Kant connects this latter sense of ‘possibility’ with reason rather than the understanding (A 232 / B 284–85).

6. This point is emphasized by Wood, *Kant's Rational Theology*, p. 59, in response to the claims of Kemp Smith and England as remnants of Kant's pre-critical views. (See note 1, above.)
7. This is noted by Wood (*Kant's Rational Theology*, p. 41), who also correctly points out that the difference between Leibniz and Kant on this point is that for the former it is an infinite task of analysis and for the latter one of synthesis.
8. For Wood's discussion of the metaphor of an "ontological space" see *Kant's Rational Theology*, esp. pp. 33–34.
9. See ND 1: 395–96; 15–17; EMB 2: 77–81; 122–26.
10. See J. F. Courtine, "Realitas," and "Realität / Idealität," in *Historisches Wörterbuch der Philosophie*, vol. 8, pp. 178–93.
11. This is noted by Wood, *Kant's Rational Theology*, pp. 31–32. Here Wood is responding to Kemp Smith's contention that Kant's appeal to this bit of traditional ontology in the Ideal is incompatible with the critical doctrines of the Analytic. See Kemp Smith, *Commentary*, pp. 524–25.
12. As Wood also points out (*Kant's Rational Theology*, pp. 31–33), there appears to be some ambiguity in Kant's conception of reality, which we need not consider here.
13. Kant explicitly distinguishes *realitas phenomenon* and *realitas noumenon* in this way in ML<sub>2</sub> 28: 560.
14. See Wood, *Kant's Rational Theology*, p. 33.
15. This *Reflexion* is cited and discussed by Longuenesse, *Kant and the Capacity to Judge*, p. 309.
16. In the *Critique* Kant's partial endorsement of, but critical distance from, the rationalist position is most clearly expressed in his discussion of the matter-form relation in the Amphiboly. As he notes there, from the point of view of logic, matter is prior to form in the sense in which concepts are prior to the judgments in which they are related. More important, the same priority holds from the ontological point of view (here Kant explicitly has Leibniz in mind), since "unbounded reality," regarded as the "matter of all possibility," must be conceived as the condition of the conceptual determination of things, which is said to proceed (as in the Ideal) by the introduction of limitations (A266–56 / B322). But when sensible intuition is brought into the picture, the relation is reversed, since the possibility of things (as objects of possible experience) presupposes a formal intuition (A267–68 / B323–24). In spite of this, however, in the Ideal Kant affirms the conceptual priority of the matter of empirical possibility (with its spatiotemporal form) over its form. For a discussion of the connections between the Amphiboly and the Ideal see Longuenesse, *Kant and the Capacity to Judge*, esp. pp. 154–56.
17. See A323 / B379 and A335 / B392.
18. This is reminiscent of Kant's claim of an analogy between category and schema in the Analytic (A181 / B224), which was discussed in chapter 8.
19. Wood, *Kant's Rational Theology*, p. 53.
20. *Ibid.*, pp. 53–54.
21. We have seen that in the Analytic Kant affirms such isomorphism between logical function and category, on the one hand, and category and schema, on the other. Another important isomorphism underlies Kant's account of beauty as a symbol of the morally good

- in the *Critique of Judgment* (§59). For my discussion of the latter, which I characterize as a “reflective isomorphism,” see Allison, *Kant’s Theory of Taste*, pp. 254–63.
22. See Kemp Smith, *Commentary*, pp. 524–25; England, *Kant’s Conception of God*, p. 120. For a fuller response to this line of criticism, see Wood, *Kant’s Rational Theology*, pp. 56–59.
  23. See also A 579 / B 607 and A 580 / B 608. In the latter place he characterizes the concept of an *ens realissimum* as “a mere fiction, through which we encompass and realize the manifold of our idea in an ideal, as a particular being.”
  24. Once again, I am largely following Grier, *Kant’s Doctrine of Transcendental Illusion*, chapter 7. This provides what I take to be the definitive account of the roles of transcendental illusion and metaphysical error in Kant’s critique of rational theology.
  25. See chapter II, note 5.
  26. This identification was suggested to me by Grier in correspondence.
  27. In Kant’s terms, both are *conceptus ratiocinantes* as opposed to *conceptus ratiocinati* (A 311 / B 368).
  28. The distinction is also referred to, though not explained, at A 615 / B 643.
  29. On this point see Caimi, “On a Non-Regulative Function of the Ideal of Reason,” *Proceedings of the Eighth International Kant Congress*, pp. 540–41.
  30. This recalls the discussion of ‘absolute’ at A 324–26 / B 380–82. It also differs from the sense of necessity at issue in the Postulates and in this respect parallels the notion of absolute possibility discussed earlier (see note 5).
  31. The issue is further complicated by the fact that Kant likewise provides a version of a cosmological argument for an absolutely necessary being in the thesis of the Fourth Antinomy (A 452–54 / B 480–82). In the Antinomy, however, this absolutely necessary being is located in the world, either as part or cause, which suggests a Spinozistic conception of God, whereas the Ideal is explicitly concerned with a transcendent being.
  32. Schopenhauer, *The World as Will and Representation*, vol. I, p. 511.
  33. See ND I: 395–96; 15–17, and EMB 2: 72–73; 117–19.
  34. The Ideal also contains a critique of the physico-theological proof (A 620–30 / B 648–58), which is the third of three possible speculative proofs recognized by Kant. But since this does not belong to transcendental theology (because of its strong empirical premise), yet (like the cosmological proof) depends crucially on the ontological, it need not be considered here.
  35. Here as elsewhere, Kant distinguishes between the possibility of a concept (logical possibility), which is simply lack of contradiction, and the real possibility of a thing. But he does not press the issue at this point (see A 596 / B 624 note).
  36. This line of objection is raised by Shaffer, “Existence, Predication, and the Ontological Argument,” in *The First Critique, Reflections on Kant’s Critique of Pure Reason*, p. 125. It is repeated but not endorsed by Wood, *Kant’s Rational Theology*, pp. 105–07, whose response to it is similar to my own.
  37. See Shaffer, “Existence, Predication, and the Ontological Argument,” p. 126, and Wood, *Kant’s Rational Theology*, pp. 108–09. Both Shaffer and Wood express astonishment that this argument, which they obviously view as extremely weak, has stood up so well under scrutiny.

38. Kant identifies it with the Leibnizian proof from the contingency of the world (A604 / B 632).
39. In this respect it differs sharply from the physico-theological proof, which argues from the determinate experience of order in the world to the existence of an intelligent author. Among the problems with this argument, however, is that it cannot get us to the *ens realissimum*. Consequently, at the end of the day it too must fall back on the ontological argument.
40. See A615–16 / B 643–44. Although Kant does not refer explicitly to  $P_2$  here, he clearly has it in mind.
41. For an account of the relation between illusion and fallacy in the ontological proof, see Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 256–60.
42. See, for example, Dennett, “Mechanism and Responsibility,” in *Essays on Freedom of Action*, ed. by Ted Honderich, pp. 157–84; and *The Intentional Stance*.
43. Perhaps to avoid this misleading implication, the phrase is often translated as “principles of harmony.”
44. A good example of this way of thinking is provided by Wood, *Kant's Rational Theology*, pp. 76–78. Focusing on the distinction between “subjective” and “objective validity,” rather than necessity, Wood questions whether Kant's attribution of merely the former to the Ideal is defensible. And in support of this he appeals to the scientific notion of the “best explanation,” which is usually taken as evidence for the objective validity of an hypothesis. The problem is that a Kantian idea is not a putative explanatory hypothesis—though there seems no other way for the transcendental realist, who desires to treat it as something more than a useful fiction, to regard it.

## CHAPTER 15

1. Kant further contributes to this view by suggesting in the *Prolegomena* that an essential portion of the Appendix has the character of a “scholium” to the discussion of metaphysics (Pro 4: 364; 118).
2. The classical statement of this reading is by Kemp Smith, *Commentary*, pp. 543–52.
3. See also A654 / B 682.
4. This is particularly insisted upon by Rolf-Peter Horstmann, “Why Must There Be a Transcendental Deduction in Kant's *Critique of Judgment*?,” in *Kant's Transcendental Deductions, The Three “Critiques” and the “Opus Postumum,”* ed. by Eckart Förster, pp. 157–76, and “Der Anhang zur transzendentalen Dialektik,” in *Kritik der reinen Vernunft*, ed. by Georg Mohr and Marcus Willaschek, pp. 525–45.
5. These include, but are by no means limited to, Buchdahl, *Metaphysics and the Philosophy of Science*, esp. pp. 641–81, and *Kant and the Dynamics of Reason*; Butts, *Kant and the Double Government Methodology*; Kitcher, “Projecting the Order of Nature,” in *Kant's Philosophy of Science*, ed. by Butts, pp. 201–35, reprinted in *Kant's Critique of Pure Reason, Critical Essays*, ed. by Patricia Kitcher, pp. 219–38; McFarland, *Kant's Concept of Teleology*, esp. pp. 1–42; Rescher, *Kant and the Reach of Reason*; and Wartenberg, “Reason and the Practice of Science,” in *The Cambridge Companion to Kant*, pp. 228–48.

6. The point is forcefully argued by Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 263–305. Thus, my account is once again greatly indebted to hers. Although a connection between Kant's doctrine of illusion and the regulative function of ideas is also noted by other commentators, it is not developed in a way that really helps to illuminate Kant's views on the epistemic function of reason. These include Heimsoeth, *Transzendente Dialektik*, vol. 3, pp. 554–55; Piché, *Das Ideal: Ein Problem der Kantischen Ideenlehre*, pp. 36–43; and Rescher, *Kant and the Reach of Reason*, pp. 152–55.
7. See Grier, *Kant's Doctrine of Transcendental Illusion*, pp. 37–38.
8. The text of Meier's *Auszug* is reprinted in KGS, vol. 16. This volume contains Kant's *Reflexionen* dealing with logic, which are largely based on this work. (See chapter 6, note 23.)
9. Meier, *Auszug* §401, KGS 16: 752–53.
10. Meier, *Vernunftlehre* §429. The passage is cited by Adickes, KGS 16, p. 753.
11. Hume, *An Enquiry Concerning Human Understanding*, p. 16. Although both Meier and Kant leave out memory, this does not materially affect the issue.
12. The *Reflexionen* are 3275–3294 16: 753–61. According to Kant, who is following Meier here, induction and analogy are the two ways of inferring (empirical) universals from particulars. The former moves from the particular to the universal according to what Kant terms the “principle of universalization” (*Princip der Allgemeinmachung*): “What belongs to many things of a genus belongs to the remaining ones too.” The latter moves from a similarity between two things with respect to a particular property to a total similarity according to the corresponding “principle of specification: Things of one genus, which we know to agree in much, also agree in what remains, with which we are familiar in some things of this genus but which we do not perceive in others” (JL 9: 133; 626). There are similar accounts in the other versions of the lectures as well as in the *Reflexionen*.
13. See, for example, LB 24: 287; 232; LH 24: 109–10; LD-W 24: 777; 508.
14. See R3276 16: 755; R3278 16: 755–56; R3294 16: 761; JL 9: 133; 627.
15. I discuss Kant's treatment of the problem in the third *Critique*, where it is intertwined with his account of reflective judgment, in *Kant's Theory of Taste*, pp. 35–42.
16. Returning to the contemporary idiom introduced in chapter 2, it might be claimed that Kant is concerned here with the warrant in “warranted assertibility.” The thesis that Kant advances a theory of warranted assertibility, understood as a third way between scientific realism and instrumentalism, is developed by Philip Kitcher in “Projecting the Order of Nature.” This reconstruction is endorsed by both Guyer, “Kant's Conception of Empirical Law,” p. 242, and Walker, “Kant's Conception of Empirical Law—II,” p. 248. Although I have no objection to this, my own view is that the distinctiveness of Kant's position on the issue is less anachronistically expressed by viewing it as a third way between Leibnizian metaphysical realism and Lockean conventionalism. I argue for this in connection with the third *Critique* in “The *Critique of Judgment* as the ‘True Apology for Leibniz,’” in *Kant und die Berliner Aufklärung*, edited by Volker Gerhardt, Rolf-Peter Horstmann, and Ralph Schumacher, vol. 1, pp. 286–99.
17. See Walker, “Kant's Conception of Empirical Law—II,” p. 245.
18. As Michael Friedman has pointed out by appealing to the example of Kant's treatment of

chemistry (MAN 4: 468–69; 184), Kant does not hold that systematic unity is sufficient for grounding genuine laws, that is, it is not itself a “sufficient mark of empirical truth.” See Friedman, “Causal Laws and the Foundations of Natural Science,” in *The Cambridge Companion to Kant*, p. 189. This, as well as a similar claim about psychology (to be discussed below), seems to rule out viewing Kant as a sheer coherentist. It does not rule out, however, a conflict with the claim that the presupposition of a synthetic unity is a necessary condition of the search for empirical laws, which entails that without this presupposition there could be no sufficient mark of such truth.

19. See KU 5: 180; 67.
20. Similar points are made from an explicitly pragmatic perspective by Rescher, *Kant and the Reach of Reason*, pp. 75–78 and passim, and Wartenberg, “Reason and the Practice of Science,” p. 246.
21. Of course, the latter is not strictly true, since we have seen that the dynamical principles of the understanding are in one sense regulative and in another constitutive. This cannot be used to rescue Kant from the present difficulty, however, since he clearly does assume a strict regulative-constitutive distinction at the level of reason.
22. This is the central feature of Horstmann’s dismissive treatment of Kant’s claims in the Appendix. (See note 4, above.)
23. Here again, I am closely following Grier, *Kant’s Doctrine of Transcendental Illusion*, pp. 269–79.
24. The contrary is suggested by Guyer, “Reason and Reflective Judgment, Kant and the Significance of Systematicity,” *Noûs* 24 (1990), pp. 17–43.
25. See Ginsborg, *The Role of Taste in Kant’s Theory of Cognition*, p. 190.
26. I have argued elsewhere that such a hierarchical organization is a necessary condition of empirical concepts, since any such concept must have further species falling under it (its extension) and higher concepts within it (its intension). See Allison, *Kant’s Theory of Taste*, pp. 33–34.
27. He also suggests that “some [*einige*] objective validity” is required for their regulative use (A664 / B692).
28. The treatment of the problem by Rescher, with which I am broadly sympathetic, repeatedly emphasizes the distinction between objective validity and objective reality (*Kant and the Reach of Reason*, pp. 131, 137, 157–67, passim). Principles of reason, on his view, possess the former but not the latter. It should be noted, however, that at least at one point Kant attributes to them a kind of objective reality (A665 / B694). As I have tried to argue, the key lies in the nature of the “object.”
29. It should be kept in mind here that Kant characterized  $P_2$  as synthetic *a priori*. See A308 / B364.
30. The key notion here is that of “heautonomy,” which Kant introduces in the third *Critique* in order to characterize the legislation of judgment in its reflective capacity (see FI 20: 225; 28 and especially KU 5: 185–86; 72). By characterizing reflective judgment as “heautonomous” in both its cognitive and its aesthetic functions, Kant is able to underscore the point that it legislates to itself rather than either to nature or to reason. Thus, the validity of this legislation remains merely “subjective.” For my discussion of Kant’s



- use of this notion, which appears only in the Introductions, see Allison, *Kant's Theory of Taste*, esp. pp. 40–42, 45.
31. See also A678 / B705.
  32. It is not clear whether the idea of a supreme intelligence is to be distinguished from that of a full-fledged *ens realissimum*. Kant's introduction of it here as a mere illustration, reserving the actual deduction of the idea of God for later, suggests that it is to be so distinguished. Moreover, in the *Ideal* Kant distinguished between the creator of the world and a mere architect (which might be equated with a supreme intelligence) (A627 / B655); but he does not appear to be consistent on this point.
  33. In the first *Critique* Kant contrasts systematicity and purposiveness, viewing the latter in explicitly teleological terms, thereby rendering it in effect supervenient on the former. In the third *Critique*, however, Kant treats systematicity under the label of a logical or formal purposiveness, as a species of purposiveness. This reflects Kant's change in perspective from that of reason to reflective judgment.
  34. See especially A687–94 / B715–22.
  35. For example, Bennett, *Kant's Dialectic*, pp. 272–73, simply denies that the idea of the soul has any regulative function.
  36. I say systematic description rather than science because Kant claims that this is the most of which psychology is capable. See MAN 4: 471; 186.
  37. This similarity has not gone unnoticed in the literature. See, for example, Brook, *Kant and the Mind*; Kitcher, *Kant's Transcendental Psychology*; and numerous papers of Ralf Meerbote, including "Kant on Freedom and the Rational and Morally Good Will," in *Self and Nature in Kant's Philosophy*, ed. by Allen Wood, pp. 57–72; "Kant on the Non-determinate Character of Human Actions," in *Causality, Freedom, and Objectivity*, ed. by William A. Harper and Ralf Meerbote, pp. 138–63; and "Wille and Willkür in Kant's Theory of Action," in *Interpreting Kant*, ed. by Moltke S. Gram, pp. 69–80.
  38. For further discussion of this issue see Allison, *Idealism and Freedom*, esp. pp. 53–66, 92–106.
  39. Although strictly speaking there is a single cosmological idea (that of the world), Kant frequently refers to them in the plural, with a different idea underlying each of the Antinomies. We shall see below that it is convenient to refer to distinct cosmological ideas in order to distinguish among a variety of regulative functions. But this is compatible with regarding these ideas as distinct expressions of the idea of the world, just as the Antinomy is expressed in four distinct Antinomies.
  40. Kant does attempt a deduction of freedom from the "practical point of view" in both the *Groundwork* and the *Critique of Practical Reason*. For my analyses of these, see Allison, *Kant's Theory of Freedom*, pp. 214–49, and "We Can Act Only Under the Idea of Freedom."
  41. For a fuller discussion of this, see Allison, "The Antinomy of Pure Reason, Section 9," in *Kritik der reinen Vernunft*, ed. by Georg Mohr and Marcus Willaschek, esp. pp. 469–74.
  42. Here Kant is challenging the traditional view that in order for something to be infinitely divisible it must be composed of an infinite number of parts. This principle was appealed to by Hume in his critique of the thesis of the infinite divisibility of space and time (*Trea-*

*tise*, p. 23) as well as by many other thinkers. From a Kantian standpoint, this may be seen as the inevitable consequence of a transcendentially realistic approach to the question of the divisibility of a continuum.

43. In the language of the Antinomy chapter, the point can be made by noting that the transcendental realist cannot draw a principled distinction between something being *gegeben* and something being *aufgegeben*.
44. Walker, “Kant’s Conception of Empirical Law,” pp. 255–56.
45. *Ibid.*, pp. 56–57.

## BIBLIOGRAPHY

- Adams, Robert Merrihew. "Things in Themselves." *Philosophy and Phenomenological Research* 57 (1997), pp. 801–25.
- Adickes, Erich. *Kants Lehre von der doppelten Affektion unseres Ich als Schlüssel zu seiner Erkenntnistheorie*. Tübingen: J. C. Mohr, 1929.
- . *Kant und das Ding an sich*. Berlin: Pan, 1924.
- Al-Azm, Sadik J. *The Origins of Kant's Argument in the Antinomies*. Oxford: Oxford University Press, 1972.
- Allison, Henry E. "The Antinomy of Pure Reason, Section 9," in *Immanuel Kant: Kritik der reinen Vernunft*, ed. by Mohr and Willaschek (see below), pp. 465–90.
- . *Spinoza: An Introduction*. New Haven and London: Yale University Press, 1987.
- . "The Critique of Judgment as a 'True Apology' for Leibniz," in *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses*, ed. by V. Gerhardt, R.-P. Horstmann, and R. Schumacher (see below), vol. 1, pp. 286–99.
- . "The Critique of Pure Reason as Transcendental Phenomenology," in *Dialogues in Phenomenology*, ed. by Don Ihde and Richard N. Zaner, pp. 136–55. The Hague: Martinus Nijhoff, 1975.
- . *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge: Cambridge University Press, 1996.

- . "Kant and the Claims of Knowledge" (review). *Journal of Philosophy* 86 (1989), pp. 214–21.
- . "Kant's Concept of the Transcendental Object." *Kant-Studien* 59 (1968), pp. 165–86.
- . "Kant's Critique of Berkeley." *Journal of the History of Philosophy* 11 (1973), pp. 43–63.
- . "Kant's Critique of Spinoza," in *The Philosophy of Baruch Spinoza*, ed. by Richard Kennington, pp. 199–227. Washington, D.C.: Catholic University Press of America, 1980. (*Studies in Philosophy and the History of Philosophy*, vol. 7.)
- . "Kant's Non-Sequitur." *Kant-Studien* 62 (1971), pp. 367–77.
- . "Kant's Refutation of Realism," *Dialectica* 30 (1976), pp. 223–53.
- . *Kant's Theory of Freedom*. Cambridge: Cambridge University Press, 1990.
- . *Kant's Theory of Taste: A Reading of the Critique of Aesthetic Judgment*. Cambridge: Cambridge University Press, 2001.
- . *Kant's Transcendental Idealism: An Interpretation and Defense*. New Haven: Yale University Press, 1983.
- . "The Non-Spatiality of Things in Themselves for Kant," *Journal of the History of Philosophy* 14 (1976), pp. 313–21.
- . "Things in Themselves, Noumena, and the Transcendental Object," *Dialectica* 32 (1978), pp. 41–76.
- . "Transcendental Affinity—Kant's Answer to Hume?" in *Kant's Theory Of Knowledge*, ed. by L. W. Beck (see below), pp. 119–27.
- . "Transcendental Idealism and Descriptive Metaphysics," *Kant-Studien* 60 (1969), pp. 216–23.
- . "Transcendental Schematism and the Problem of the Synthetic A Priori," *Dialectica* 35 (1981), pp. 57–83.
- . "Where Have All the Categories Gone? Reflections on Longuenesse's Reading of Kant's Transcendental Deduction," *Inquiry—Interdisciplinary Journal of Philosophy* 43 (2000), pp. 57–80.
- , ed. and trans. *The Kant-Eberhard Controversy*. Baltimore: Johns Hopkins University Press, 1973.
- Ameriks, Karl. "Kant, Fichte, and Apperception," in *Kant and the Fate of Autonomy*, ed. by Karl Ameriks, pp. 234–64. Cambridge: Cambridge University Press, 2000.
- . "Kantian Idealism Today." *History of Philosophy Quarterly* 9 (1992), pp. 329–42.
- . *Kant's Theory of Mind*. Oxford: Clarendon Press, 1982.
- . "Recent Work on Kant's Theoretical Philosophy." *American Philosophical Quarterly* 19 (1982), pp. 1–24.
- Aquila, Richard E. "Personal Identity and Kant's Refutation of Idealism." *Kant-Studien* 70 (1979), pp. 257–78.
- Barker, Stephen. "Appearing and Appearances," reprinted in *Kant Studies Today*, ed. by L. W. Beck (see below), pp. 278–89.
- Baum, Manfred. "Kant on Pure Intuition," in *Minds, Ideas, and Objects*, ed. by Cummins and Zoeller (see below), pp. 303–15.

- Bayle, Pierre. *Historical and Critical Dictionary*, selections ed. and trans. by Richard H. Popkin. Indianapolis: Bobbs-Merrill, 1965.
- Beck, Jacob Sigismund. *Einzig möglicher Standpunct aus welchem die critische Philosophie beurtheilt werden muss*. Riga: Johann Friedrich Hartnoch, 1796. Reprinted in *Aetas Kantiana*. Brussels: Culture et Civilisation, 1969.
- Beck, Lewis White, "Can Kant's Synthetic Judgments Be Made Analytic?," reprinted in *Kant: Disputed Questions*, ed. by Moltke S. Gram (see below), pp. 228–46.
- . *A Commentary on Kant's "Critique of Practical Reason."* Chicago: University of Chicago Press, 1960.
- . *Essays on Kant and Hume*. New Haven: Yale University Press, 1978.
- . "Kant's Theory of Definition," reprinted in *Kant: Disputed Questions*, ed. by Moltke S. Gram (see below), pp. 215–27.
- . "Lewis' Kantianism." *Studies in the Philosophy of Kant* (see below), pp. 108–24.
- . "A Non-Sequitur of Numbing Grossness?," in *Essays on Kant and Hume* (see above), pp. 147–53.
- . "Once More into the Breach: Kant's Answer to Hume, Again," in *Essays on Kant and Hume* (see above), pp. 130–35.
- . "A Prussian Hume and a Scottish Kant," in *Essays on Kant and Hume* (see above), pp. III–29.
- . "A Reading of the Third Paragraph in B," in *Essays on Kant and Hume* (see above), pp. 141–46.
- . *Studies in the Philosophy of Kant*. Indianapolis: Bobbs-Merrill, 1965.
- , ed. *Kant's Theory of Knowledge*. Dordrecht: Reidel, 1974.
- , ed. *Kant Studies Today*. La Salle, Ill.: Open Court, 1969.
- Beiser, Frederick C. *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge, Mass.: Harvard University Press, 1987.
- Bell, David. "The Art of Judgment." *Mind* 96 (1987), pp. 221–44.
- Bennett, Jonathan. *Kant's Analytic*. Cambridge: Cambridge University Press, 1966.
- . *Kant's Dialectic*. Cambridge: Cambridge University Press, 1974.
- . *Locke, Berkeley, Hume*. Oxford: Oxford University Press, 1971.
- Berkeley, George. *The Works of George Berkeley, Bishop of Cloyne*, ed. by A. A. Luce and T. E. Jessop. London: Thomas Nelson and Sons, 1948.
- Bird, Graham. *Kant's Theory of Knowledge*. London: Routledge and Kegan Paul, 1962.
- . "Recent Interpretations of Kant's Transcendental Deduction," in *Akten des 4. Internationalen Kant-Kongress*, ed. by G. Funke and J. Kopper, part 1, pp. 1–14. Berlin: de Gruyter, 1974.
- Brandt, Reinhardt. *The Table of Judgments: Critique of Pure Reason A 67–76; B 92–101*, trans. by Eric Watkins. Atascadero, Calif.: Ridgeview Publishing, 1995. (*North American Kant Society Studies in Philosophy*, vol. 4.)
- Brittan, Gordon G. Jr. *Kant's Theory of Science*. Princeton: Princeton University Press, 1978.
- Broad, C. D. "Kant's First and Second Analogies of Experience." *Proceedings of the Aristotelian Society* 25 (1926), pp. 189–210.
- Brook, Andrew. *Kant and the Mind*. Cambridge: Cambridge University Press, 1994.

- Brouillet, R. "Dieter Henrich et 'The Proof-Structure of Kant's Transcendental Deduction': Reflexions critiques." *Dialogue* 14 (1975), pp. 639–648.
- Buchdahl, Gerd. "The Conception of Lawlikeness in Kant's Philosophy of Science," in *Kant's Theory of Knowledge*, ed. by L. W. Beck (see above), pp. 128–50.
- . *Kant and the Dynamics of Reason: Essays on the Structure of Kant's Philosophy*. Oxford: Basil Blackwell, 1992.
- . "The Kantian 'Dynamic of Reason' with Special Reference to the Place of Causality in Kant's System," in *Kant Studies Today*, ed. L. W. Beck (see above), pp. 187–208.
- . *Metaphysics and the Philosophy of Science*. Cambridge, Mass.: MIT Press, 1969.
- Buroker, Jill. *Space and Congruence: The Origins of Kant's Idealism*. Dordrecht: Reidel, 1981.
- Butts, Robert E. *Kant and the Double Government Methodology*. Amsterdam: Kluwer, 1987.
- . "Kant's Schemata as Semantical Rules," in *Kant Studies Today*, ed. L. W. Beck (see above), pp. 290–300.
- , ed. *Kant's Philosophy of Physical Science*. Dordrecht: Reidel, 1986. (University of Western Ontario Series in Philosophy and Science.)
- Caird, Edward. *The Critical Philosophy of Kant*. Glasgow: J. Maclehose, 1909.
- Caranti, Luigi. *The Development of Kant's Refutation of Idealism*, unpublished dissertation, Boston University, 2001.
- Carl, Wolfgang. "Apperception and Spontaneity." *International Journal of Philosophical Studies* 5 (1997), pp. 147–63.
- . "Die B-Deduktion," in *Immanuel Kant: Kritik der reinen Vernunft*, ed. by Mohr and Willaschek (see below), pp. 189–216.
- . *Die transzendente Deduktion der Kategorien in der 1. Auflage der Kritik der reinen Vernunft: Ein Kommentar*. Frankfurt: Vittorio Klostermann, 1992.
- Chipman, Lachlan. "Kant's Categories and Their Schematism." *Kant-Studien* 63 (1972), pp. 36–49.
- Ciami, Mario. "On a Non-Regulative Function of the Ideal of Pure Reason," in *Proceedings of the Eighth International Kant Congress*, ed. by Hoke Robinson, vol. 1, part 2, pp. 539–49. Milwaukee: Marquette University Press, 1995.
- Cohen, Hermann. *Kants Theorie der Erfahrung*. Berlin: Dimmler, 1871 (2nd ed., 1885).
- . *Kommentar zu Immanuel Kants Kritik der reinen Vernunft*. 4th ed. Leipzig: Felix Meiner, 1925.
- Cohn, Jonas. *Geschichte des Unendlichkeitsproblems in abendländischen Denken bis Kant*. Hildesheim: Georg Olms, 1960.
- Cummins, Phillip D., and Zoeller, Guenter, eds. *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*. Atascadero, Calif.: Ridgeview, 1992. (North American Kant Society Studies in Philosophy, vol. 2.)
- Curtius, Ernst Robert. "Das Schematismuskapitel in der Kritik der reinen Vernunft." *Kant-Studien* 19 (1916), pp. 338–66.
- Dennett, Daniel. *The Intentional Stance*. Cambridge, Mass., and London: MIT Press, 1987.
- . "Mechanism and Responsibility," in *Essays on Freedom of Action*, ed. by Ted Honderich, pp. 157–84. London: Routledge and Kegan Paul, 1973.

- Descartes, René. *The Philosophical Writings of Descartes*, ed. by J. Cottingham, R. Stoothoff, and D. Murdoch. 2 vols. Cambridge: Cambridge University Press, 1985.
- De Vleeschauwer, H. J. *La déduction transcendentale dans l'oeuvre de Kant*. 3 vols. Paris: Leroux, 1934–37.
- Donnellan, Keith S. "Reference and Definite Descriptions." *Philosophical Review* 75 (1968), pp. 281–304.
- Dryer, D. P. *Kant's Solution for Verification in Metaphysics*. Toronto: University of Toronto Press, 1966.
- Ebbinghaus, Julius. "Kants Lehre von der Anschauung a priori," in *Kant: Zur Deutung seiner Theorie von Erkennen und Handeln*, ed. by G. Prauss (see below), pp. 44–61.
- Eberhard, J. A., ed. *Philosophisches Archiv*. 2 vols. Berlin: C. Massdorf, 1792–93. Reprinted in *Aetas Kantiana*. Brussels: Culture et Civilisation, 1968.
- . *Philosophisches Magazin*. 4 vols. Halle: J. J. Gebouer, 1789–92. Reprinted in *Aetas Kantiana*. Brussels: Culture et Civilisation, 1968.
- Engel, Morris S. "Kant's Copernican Analogy: A Re-examination." *Kant-Studien* 54 (1963), pp. 243–51.
- England, F. E. *Kant's Conception of God*. New York: Humanities Press, 1968.
- Erdmann, Benno. *Kants Kriticismus in der ersten und in der zweiten Auflage der Kritik der reinen Vernunft*. Leipzig: Leopold Voss, 1878.
- Evans, Claude. "Two-Steps-in-One-Proof: The Structure of the Transcendental Deduction." *Journal of the History of Philosophy* 28 (1990), pp. 553–70.
- Ewing, A. C. *Kant's Treatment of Causality*. London: Kegan Paul, Trench, Trubner and Co., 1924.
- . *A Short Commentary on Kant's "Critique of Pure Reason."* 2d ed. London: Methuen, 1950.
- Falkenstein, Lorne. "Hume on Manners of Disposition and the Ideas of Space and Time." *Archiv für Geschichte der Philosophie* 79 (1997), pp. 179–201.
- . "Kant's Argument for the Non-spatiotemporality of Things in Themselves." *Kant-Studien* 80 (1989), pp. 265–83.
- . *Kant's Intuitionism: A Commentary on the Transcendental Aesthetic*. Toronto: University of Toronto Press, 1995.
- Ferrarin, Alfredo. "Kant on the Exhibition of a Concept in Intuition." *Kant-Studien* 86 (1995), pp. 131–74.
- Fichte, J. G. *Zweite Einleitung in die Wissenschaftslehre*, in *Erste and zweite Einleitung in die Wissenschaftslehre*, ed. by Fritz Medicus. Hamburg: Felix Meiner, 1961.
- Franzwa, Gregg E. "Space and the Schematism." *Kant-Studien* 69 (1978), pp. 149–59.
- Freudiger, Jürg. "Zum Problem der Wahrnehmungsurteile in Kants theoretische Philosophie." *Kant-Studien* 82 (1991), pp. 414–35.
- Friedman, Michael. *Kant and the Exact Sciences*. Cambridge, Mass.: Harvard University Press, 1992.
- . "Kant on Causal Laws and the Foundations of Natural Science," in *The Cambridge Companion to Kant*, ed. by Paul Guyer (see below), pp. 161–97.
- Gerhardt, Volker, Rolf-Peter Horstmann, and Ralph Schumacher, eds. *Kant und die Berliner*

- Aufklärung: Akten des IX. Internationalen Kant-Kongresses*. Berlin, New York: Walter de Gruyter, 2001.
- Gibbons, Sarah. *Kant's Theory of Imagination: Bridging Gaps in Judgment and Experience*. Oxford: Oxford University Press, 1994.
- Gibson, James. *Locke's Theory of Knowledge and Its Historical Relations*. Cambridge: Cambridge University Press, 1917.
- Ginsborg, Hannah. "Lawfulness Without a Law: Kant on the Free Play of Imagination and Understanding." *Philosophical Topics* 25 (1997), pp. 37–81.
- . *The Role of Taste in Kant's Theory of Cognition*. New York: Garland Publishing, 1990.
- Glouberman, M. "Conceptuality: An Essay in Retrieval." *Kant-Studien* 70 (1979), pp. 383–408.
- Gochnauer, Myron. "Kant's Refutation of Idealism." *Journal of the History of Philosophy* 12 (1974), pp. 195–206.
- Gram, Moltke S. "The Crisis of Synthetcity: The Kant-Eberhard Controversy." *Kant-Studien* 7 (1980), pp. 155–80.
- . "How to Dispense with Things in Themselves (1)." *Ratio* 18 (1976), pp. 1–15.
- . "Intellectual Intuition: The Continuity Thesis." *Journal of the History of Ideas* 42 (1981), pp. 287–334.
- . "Kant's First Antinomy." *Monist* 51 (1967), pp. 499–518.
- . *Kant, Ontology, and the A Priori*. Evanston, Ill.: Northwestern University Press, 1968.
- . "The Myth of the Double Affection," in *Reflections on Kant's Philosophy*, ed. by W. H. Werkmeister (see below), pp. 29–69.
- . *The Transcendental Turn*. Gainesville: University of Florida Press, 1984.
- . ed. *Kant: Disputed Questions*. Chicago: Quadrangle Books, 1967.
- Grier, Michelle. *Kant's Doctrine of Transcendental Illusion*. Cambridge: Cambridge University Press, 2001.
- Gurwitsch, Aron. *Leibniz: Philosophie des Panlogismus*. Berlin: de Gruyter, 1974.
- Guyer, Paul, ed. *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press, 1992.
- . *Kant and the Claims of Knowledge*. Cambridge and New York: Cambridge University Press, 1987.
- . "Kant's Conception of Empirical Law." *Proceedings of the Aristotelian Society*, supp. vol. (1990), pp. 220–42.
- . "Kant's Intentions in the Refutation of Idealism." *Philosophical Review* 92 (1983), pp. 329–83.
- . "Kant's Tactics in the Transcendental Deduction," in *Essays on Kant's Critique of Pure Reason*, ed. by J. N. Mohanty and Robert W. Shahan, pp. 157–99. Norman, Okla.: University of Oklahoma Press, 1982.
- . "The Postulates of Empirical Thinking in General and the Refutation of Idealism," in *Kant: Kritik der reinen Vernunft*, ed. by Morh and Willaschek (see below), pp. 297–324.
- . "Reason and Reflective Judgment: Kant on the Significance of Systematicity." *Noûs* 24 (1990), pp. 17–43.



- Hanson, Norwood Russell. "Copernicus' Role in Kant's Revolution." *Journal of the History of Ideas* 20 (1959), pp. 274–81.
- Harper, William L., and Ralph Meerbote, eds. *Kant on Causality, Freedom, and Objectivity*. Minneapolis: University of Minnesota Press, 1984.
- Heidemann, D. H. *Kant und das Problem des metaphysischen Idealismus*. Berlin and New York: de Gruyter, 1998.
- Heidemann, Ingeborg. *Spontaneität und Zeitlichkeit*. Cologne: Kölner Universitäts-Verlag, 1958.
- Heimsoeth, Heinz. "Personlichkeitsbewusstsein und Ding an sich in der kantischen Philosophie," in *Studien zur Philosophie Immanuel Kants: Metaphysische Ursprung und Ontologische Grundlagen*, pp. 227–257. Cologne: Kölner Universitäts-Verlag, 1956.
- . *Transzendente Dialektik*. 4 vols. Berlin: de Gruyter, 1969.
- Henrich, Dieter. "Die Beweisstruktur der transzendentalen Deduktion der reinen Verstandesbegriffe—eine Diskussion mit Dieter Henrich," in *Probleme der "Kritik der reinen Vernunft"*, ed. by Burkhard Tuschling. Berlin and New York: Walter de Gruyter, 1984.
- . "Fichte's ursprüngliche Einsicht," in *Subjektivität und Metaphysik*, pp. 188–232. Frankfurt: Klostermann, 1967.
- . "Fichte's Original Insight," trans. by David Lachterman, in *Contemporary German Philosophy*, vol. 1, pp. 15–53. University Park, Pa.: Pennsylvania State University Press, 1982.
- . *Identität und Objectivität: Eine Untersuchung über Kants transzendental Deduktion*. Heidelberg: Carl Winter Universitäts-Verlag, 1976.
- . "The Proof-Structure of Kant's Transcendental Deduction." *Review of Metaphysics* 22 (1969), pp. 640–59.
- Herring, Herbert. *Das Problem der Affektion bei Kant*. *Kant-Studien, Ergänzungshefte* 67 (1953).
- Hinske, Norbert. *Kants Weg zur Transzendentalphilosophie: Der dreissigjährige Kant*. Stuttgart and Berlin: W. Kohlhammer, 1970.
- Hintikka, Jaakko. "Kant on the Mathematical Method," in *Kant Studies Today*, ed. L. W. Beck (see above), pp. 117–40.
- . "On Kant's Notion of Intuition (*Anschauung*)," in *The First Critique*, ed. by T. Penelhum and J. MacIntosh (see below), pp. 38–53.
- Holt, Dennis C. "Timelessness and the Metaphysics of Temporal Existence." *American Philosophical Quarterly* 18 (1981), pp. 149–56.
- Horstmann, Rolf Peter. "Der Anhang zur transzendentalen Dialektik," in *Immanuel Kant: Kritik der reinen Vernunft*, ed. by Mohr and Willaschek (see below), pp. 525–45.
- . "Kants Paralogismen." *Kant-Studien* 83 (1993), pp. 408–25.
- . "The Metaphysical Deduction in Kant's Critique of Pure Reason." *Philosophical Forum* 13 (1981), pp. 32–47.
- . "Space as Intuition and Geometry." *Ratio* 18 (1976), pp. 17–30.
- . "Why Must There Be a Transcendental Deduction in Kant's Critique of Judgment?" in *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*, ed. by Eckart Förster, pp. 157–76. Stanford, Calif.: Stanford University Press, 1989.
- Hossenfelder, Malte. "Allison's Defense of Kant's Transcendental Idealism." *Inquiry* 33 (1990), pp. 467–79.

- . *Kants Konstitutions-theorie und die Transzendente Deduktion*. Berlin: de Gruyter, 1978.
- Hudson, Hud. *Kant's Compatibilism*. Ithaca and London: Cornell University Press, 1994.
- Hume, David. *An Enquiry Concerning Human Understanding*, ed. by Eric Steinberg. Indianapolis: Hackett Publishing Company, 1977.
- . *A Treatise of Human Nature*, ed. by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press, 2000.
- Humphrey, Ted. "The Historical and Conceptual Relations between Kant's Metaphysics of Space and Philosophy of Geometry." *Journal of the History of Philosophy* 11 (1973), pp. 483–512.
- Jacobi, F. H. *David Hume über den Glauben, oder Idealismus und Realismus, in Ueber den transzendentalen Idealismus*, in *Werke*, ed. by F. Roth and F. Köppen, vol. 2, pp. 291–310. Darmstadt: Wissenschaftliche Buchgesellschaft, 1968.
- Keller, Pierre. *Kant and the Demands of Self-Consciousness*. Cambridge: Cambridge University Press, 2001.
- Kemp Smith, Norman. *A Commentary to Kant's "Critique of Pure Reason"*. 2d ed., revised and enlarged. New York: Humanities Press, 1962.
- Kenny, Anthony. *Descartes: A Study of His Philosophy*. New York: Random House, 1968.
- Kerzberg, P. "Two Senses of Kant's Copernican Revolution." *Kant-Studien* 80 (1989), pp. 63–80.
- Kitcher, Patricia. *Kant's Transcendental Psychology*. New York: Oxford University Press, 1990.
- Kitcher, Philip. "Projecting the Order of Nature," in *Kant's Critique of Pure Reason: Critical Essays*, ed. by Patricia Kitcher, pp. 219–38. Lanham, Md: Rowman and Littlefield, 1998.
- Kneale, William, and Martha Kenale. *The Development of Logic*. Corrected paperback edition. Oxford: Clarendon Press, 1984. (1st ed. 1962.)
- Korner, Stephan. *Kant*. London: Penguin Books, 1955.
- Kotzin, Rhoda, and Jörg Baumgärtner. "Sensations and Judgments of Perceptions: Diagnosis and Rehabilitation of Some of Kant's Misleading Examples." *Kant-Studien* 81 (1990), pp. 401–412.
- Krausser, Peter. "'Form of Intuition' and 'Formal Intuition' in Kant's Theory of Experience and Science." *Studies in the History and Philosophy of Science* 4 (1973–74), pp. 279–87.
- . "The Operational Conception of 'Reine Anschauung' in Kant's Theory of Experience and Science." *Studies in the History and Philosophy of Science* 3 (1972–73), pp. 81–87.
- Krüger, L. "Wollte Kant die Vollständigkeit seiner Urteilstafel beweisen?" *Kant-Studien* 59 (1968), pp. 333–55.
- Kühn, Manfred. "Hume's Antinomies." *Hume Studies* 9 (1983), pp. 25–45.
- Lachize-Rey, Pierre. *L'idéalisme kantien*. Paris: Librairie Philosophique J. Vrin, 1950.
- Langton, Rae. *Kantian Humility: Our Ignorance of Things in Themselves*. Oxford: Oxford University Press, 1998.
- Lauener, Henri. *Hume und Kant*. Bern: Francke Verlag, 1969.
- Leibniz, G. W. *Discourse on Metaphysics*, trans. by G. R. Montgomery. La Salle, Ill.: Open Court, 1962.
- . *The Leibniz-Clarke Correspondence*, ed. and trans. by H. G. Alexander. Manchester: Manchester University Press, 1956.

- . *New Essays on Human Understanding*, trans. by P. Remnant and J. Bennett. Cambridge: Cambridge University Press, 1981.
- Locke, John. *An Essay Concerning Human Understanding*, ed. by A. C. Fraser. 2 vols. New York: Dover, 1959.
- Longuenesse, Beatrice. "The Divisions of the Transcendental Logic and the Leading Thread," in *Immanuel Kant: Kritik der reinen Vernunft*, ed. by Mohr and Willaschek (see below), pp. 131–58.
- . *Kant and the Capacity to Judge*, trans. by Charles Wolfe. Princeton: Princeton University Press, 1998.
- Lovejoy, Arthur O. "On Kant's Reply to Hume," in *Kant: Disputed Questions*, ed. Moltke S. Gram (see above), pp. 284–308.
- MacLachlan, D. L. C. "The Thing in Itself Appears in a Meta-Language," in *Proceedings of the Eighth International Kant Congress*, ed. by Hoke Robinson, vol. 2, pp. 155–61. Milwaukee: Marquette University Press, 1995.
- Makkreel, Rudolf A. *Imagination and Interpretation in Kant: The Hermeneutical Import of the "Critique of Judgment."* Chicago: University of Chicago Press, 1990.
- Martin, Gottfried. *Kant's Metaphysics and Theory of Science*, trans. by P. G. Lucas. Manchester: Manchester University Press, 1955.
- Matthews, H. E. "Strawson on Transcendental Idealism." *Philosophical Quarterly* 19 (1969), pp. 204–20.
- McCann, Edwin. "Skepticism and Kant's B Deduction." *History of Philosophy Quarterly* 2 (1985), pp. 71–90.
- McFarland, J. D. *Kant's Concept of Teleology*. Edinburgh: University of Edinburgh Press, 1970.
- Meerbote, Ralf. "Kant on the Nondeterminate Character of Human Actions," in *Causality, Freedom, and Objectivity*, ed. by William A. Harper and Ralf Meerbote (see above), pp. 138–63.
- . "Kant's Use of the Notions 'Objective Reality' and 'Objective Validity.'" *Kant-Studien* 63 (1972), pp. 51–58.
- . "Space and Time and Objects in Space and Time: Another Aspect of Kant's Transcendental Idealism," in *Minds, Ideas, and Objects*, ed. by Phillip D. Cummins and Guenter Zoeller (see above), vol. 2, pp. 275–90.
- . "The Unknowability of Things in Themselves," in *Kant's Theory of Knowledge*, ed. by L. W. Beck (see above), pp. 166–74.
- . "Wille and Willkür in Kant's Theory of Action", in *Interpreting Kant*, ed. Moltke S. Gram (see above), pp. 69–80.
- Melnick, Arthur. *Kant's Analogies of Experience*. Chicago: University of Chicago Press, 1973.
- Mohr, Georg, and Marcus Willaschek, eds. *Immanuel Kant: Kritik der reinen Vernunft*. Berlin: Akademie-Verlag, 1998. (*Klassiker Auslegen*, vols. 17/18.)
- Moore, G. E. *Some Main Problems of Philosophy*. London: George Allen and Unwin, 1953.
- Morrison, Margaret C. "Community and Coexistence: Kant's Third Analogy of Experience." *Kant-Studien* 89 (1998), pp. 257–77.
- Ortwein, Birger. *Kants problematische Freiheitslehre*. Bonn: Bouvier, 1983.
- O'Shaughnessy, Brian. *The Will: A Dual-Aspect Theory*. 2 vols. Cambridge: Cambridge University Press, 1980.

- Parsons, Charles. "Kant's Philosophy of Arithmetic," in *Philosophy, Science, and Method*, ed. by S. Morgenbesser, P. Suppes, and M. White, pp. 568–94. New York: St. Martin's Press, 1969.
- . "The Transcendental Aesthetic," in *The Cambridge Companion to Kant*, ed. by Paul Guyer (see above), pp. 62–100.
- Paton, H. J. *The Categorical Imperative*. London: Hutchinson, 1958.
- . *Kant's Metaphysic of Experience*. 2 vols. New York: Macmillan, 1936.
- Penelhum, T., and J. MacIntosh, eds. *The First Critique: Reflections on Kant's Critique of Pure Reason*. Belmont, Calif.: Wadsworth, 1969.
- Piché, Claude. *Das Ideal: Ein Problem der kantischen Ideenlehre*. Bonn: Bouvier, 1984.
- Pippin, Robert B. "Kant on the Spontaneity of Mind." *Canadian Journal of Philosophy* 17 (1987), pp. 449–76.
- . *Kant's Theory of Form*. New Haven and London: Yale University Press, 1982.
- Posy, Carl. "Transcendental Idealism and Causality," in *Kant on Causality*, ed. by Harper and Meerbote (see above), pp. 20–41.
- Prauss, Gerold. *Erscheinung bei Kant*. Berlin: de Gruyter, 1971.
- . *Kant und das Problem der Dinge an sich*. Bonn: Bouvier, 1974.
- , ed. *Kant: Zur Deutung seiner Theorie von Erkennen und Handeln*. Cologne: Kienpauer and Witsch, 1973.
- Prichard, H. A. *Kant's Theory of Knowledge*. Oxford: Clarendon Press, 1909.
- Putnam, Hilary. *Reason, Truth, and History*. Cambridge: Cambridge University Press, 1981.
- Reich, Klaus. *The Completeness of Kant's Table of Judgments*, with J. E. Kneeller, trans. by Jane Keller and Michael Losonsky. Stanford: Stanford University Press, 1992.
- . *Die Vollständigkeit der kantischen Urteilstafel*. Berlin: Rostock, 1932.
- Reid, Thomas. *Philosophical Works*, ed. by William Hamilton. 2 vols. Reprinted by Georg Olms Verlagsbuchhandlung, Hildesheim, 1967.
- Reinhold, Karl Leonard. *Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens*. Prague and Jena: C. Witzmann and I. M. Marke, 1789.
- Rescher, Nicholas. *Kant and the Reach of Reason: Studies in Kant's Theory of Rational Systematization*. Cambridge: Cambridge University Press, 1999.
- . "Noumenal Causality," in *Kant's Theory of Knowledge*, ed. by L. W. Beck (see above), pp. 175–83.
- Robinson, Hoke. "Intuition and Manifold in the Transcendental Deduction." *Southern Journal of Philosophy* 22 (1984), pp. 403–12.
- . "Two Perspectives on Kant's Appearances and Things in Themselves." *Journal of the History of Philosophy* 32 (1994), pp. 411–41.
- Rohs, Peter. "Kants Prinzip der durchgängigen Bestimmung alles Seienden." *Kant-Studien* 69 (1978), pp. 170–80.
- Rosas, Alejandro. *Kants idealistische Reduktion: Das Mentale und das Materielle im transzendentalen Idealismus*. Würzburg: Königshausen und Neumann, 1996.
- Röttges, Heinz. "Kants Auflösung der Freiheitsantinomie." *Kant-Studien* 65 (1974), pp. 33–49.
- Roussett, Bernard. *La doctrine kantienne de l'objectivité*. Paris: Librairie Philosophique J. Vrin, 1967.

- Russell, Bertrand. *Our Knowledge of the External World*. London: George Allen and Unwin, 1914.
- Schaper, Eva. "The Kantian Thing-in-itself as a Philosophical Fiction." *Philosophical Quarterly* 16 (1966), pp. 233–43.
- . "Kant's Schematism Reconsidered." *Review of Metaphysics* 8 (1964), pp. 267–92.
- Schopenhauer, Arthur. *Über die vierfache Wurzel des Satzes vom zureichenden Grunde*, in *Sämtliche Werke*, ed. by J. Frauenstadt. vol. 1. Leipzig: F. U. Brochhaus, 1919.
- . *The World as Will and Representation*, trans. by E. F. Payne. New York: Dover, 1969.
- Schulze, G. E. *Aenesidemus oder über die Fundamente der von dem Herrn Prof. Reinhold in Jena gelieferten Elementar-Philosophie*. 1792. Reprinted in *Aetas Kantiana*. Brussels: Culture et Civilisation, 1969.
- Schulze (or Schultz), J. G. *Erleuterungen über des Herrn Professor Kant Kritik der reinen Vernunft*. Königsberg: Hartung, 1791. Reprinted in *Aetas Kantiana*. Brussels: Culture et Civilisation, 1968.
- . *Prüfung der kantischen Kritik der reinen Vernunft*. Königsberg: Hartung, 1789. Reprinted in *Aetas Kantiana*. Brussels: Culture et Civilisation, 1968.
- Seel, Gerhard. "Die Einleitung in die Analytik der Grundsätze, der Schematismus und die obersten Grundsätze (A130/B169–A158/B197)," in *Immanuel Kant: Kritik der reinen Vernunft*, ed. by Georg Mohr and Marcus Willaschek (see above), pp. 217–46.
- Seidl, Horst. "Bemerkungen zu Ding an sich und Transzendentalen Gegenstand in Kant's Kritik der reinen Vernunft." *Kant-Studien* 63 (1972), pp. 305–14.
- Sellers, Wilfred. "The Role of the Imagination in Kant's Theory of Experience," in *Categories: A Colloquium*, ed. by Henry W. Johnstone, Jr. University Park: Pennsylvania State University, 1978.
- . *Science and Metaphysics: Variations on Kantian Themes*. London: Routledge and Kegan Paul, and New York: Humanities Press, 1968.
- . "... This I or he or it (the thing) which thinks. . . ." *Proceedings and Addresses of the American Philosophical Association* 44 (September 1971), pp. 5–31.
- Serck-Hanssen, Camilla. "Transcendental Apperception: A Study of Kant's Theoretical Philosophy and Idealism," Ph.D. dissertation, University of California, San Diego, 1996 (unpublished).
- Shaffer, J. "Existence, Predication, and the Ontological Argument." *Mind* 71 (1962), pp. 307–25.
- Shoemaker, Sydney. *Self-Knowledge and Self-Identity*. Ithaca: Cornell University Press, 1963.
- Spinoza, Benedict de. *The Chief Works of Spinoza*, trans. by R. H. M. Elwes. 2 vols. New York: Dover, 1951.
- Stenius, Erik. "On Kant's Distinction between Phenomena and Noumena," in *Philosophical Essays Dedicated to Gunnar Aspelin on the Occasion of His 65th Birthday*, pp. 231–45. Lund: C. W. R. Gleerup, 1965.
- Strawson, P. F. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. London: Methuen, 1966.
- . "Imagination and Perception," reprinted in Ralph Walker, *Kant on Pure Reason* (see below), pp. 82–99.

- Stuhlmann-Laeisz, Rainer. *Kants Logik*. Berlin: de Gruyter, 1976. (Quellen und Studien zur Philosophie, vol. 9.)
- Sturma, Dieter. *Kant über Selbstbewusstsein: Zum Zusammenhang von Erkenntniskritik und Theorie des Selbstbewusstseins*. Hildesheim, Zürich, and New York: Goerg Olms Verlag, 1985.
- Suchting, W. A. "Kant's Second Analogy of Experience." *Kant-Studien* 58 (1967), pp. 355–69.
- Svendsen, Lars Fr. H. *Kant's Critical Hermeneutics: On Schematization and Interpretation*. Dissertation. Oslo: Unipub Forlag/Akademika AS, 1999.
- Swinburne, Richard. *Time and Space*. London: Macmillan, 1968.
- Tetens, Johann Nicholas. *Philosophische Versuche über die menschliche Natur und ihre Entwicklung*. 2 vols. Leipzig: M. G. Weidmanns Erben und Reich, 1777.
- Thielke, Peter. "Discursivity and Its Discontents: Maimon's Challenge to Kant's Account of Cognition." Unpublished Ph.D. dissertation. University of California, San Diego, 1999.
- Thöle, Bernhard. "Die Analogien der Erfahrung," in *Immanuel Kant: Kritik der reinen Vernunft*, ed. by Mohr and Willaschek (see above), pp. 267–96.
- . *Kant und das Problem der Gesetzmässigkeit der Natur*. Berlin: de Gruyter, 1991. (Quellen und Studien zur Philosophie, vol. 27.)
- Thompson, Manley. "Singular Terms and Intuitions in Kant's Epistemology." *Review of Metaphysics* 26 (1972), pp. 314–43.
- Tonelli, Giorgio. "Die Voraussetzungen zur Kantischen Urteilstafel in der Logik des 18. Jahrhunderts," in *Kritik und Metaphysik: Festschrift für H. Heimsoeth*. Berlin: de Gruyter, 1966.
- Trendelenburg, Adolf. *Logische Untersuchungen*. Leipzig: Hinzel, 1862.
- Turbayne, Colin. "Kant's Refutation of Dogmatic Idealism." *Philosophical Quarterly* 5 (1955), pp. 225–44.
- Vaihinger, H. *Commentar zu Kants Kritik der reinen Vernunft*. 2 vols. Stuttgart: W. Spemann, 1881–92.
- Van Cleve, James. "The Argument from Geometry, Transcendental Idealism, and Kant's Two Worlds," in *Minds, Ideas, and Objects*, ed. by Cummins and Zoeller (see above), pp. 296–300.
- . "Four Recent Interpretations of Kant's Second Analogy." *Kant-Studien* 64 (1973), pp. 69–87.
- . *Problems from Kant*. Oxford: Oxford University Press, 1999.
- . "Substance, Matter, and Kant's First Analogy." *Kant-Studien* 70 (1979), pp. 149–61.
- Wagner, Hans. "Der Argumentationsgang in Kants Deduktion der Kategorien." *Kant-Studien* 71 (1980), pp. 352–366.
- Walker, Ralph. *Kant. The Arguments of the Philosophers*. London: Routledge and Kegan Paul, 1978.
- . "Kant's Conception of Empirical Law—II." *Aristotelian Society Supplement* 64 (1990), pp. 243–58.
- . "The Status of Kant's Theory of Matter," in *Kant's Theory of Knowledge*, ed. L. W. Beck (see above), pp. 151–56.

- , ed. *Kant on Pure Reason*. Oxford: Oxford University Press, 1982.
- Walsh, W. H. *Kant's Criticism of Metaphysics*. Edinburgh: Edinburgh University Press, 1975.
- . "Kant on the Perception of Time," in *The First Critique*, ed. by T. Penelhum and J. MacIntosh (see above), pp. 70–88.
- Warnock, G. J. "Concepts and Schematism." *Analysis* 8 (1949), pp. 77–82.
- Warren, Daniel. "Kant and the Apriority of Space," *Philosophical Review* 107 (1998), pp. 179–224.
- Wartenberg, Thomas E. "Reason and the Practice of Science," in *The Cambridge Companion to Kant*, ed. by Paul Guyer (see above), pp. 228–48.
- Washburn, Michael. "The Problem of Self-Knowledge and the Evolution of the Critical Epistemology, 1781 and 1787." Unpublished Ph.D. dissertation, University of California, San Diego, 1970.
- Watkins, Eric. "Kant's Third Analogy of Experience." *Kant-Studien* 88 (1997), pp. 406–41.
- Waxman, Wayne. *Kant's Model of the Mind: A New Interpretation of Transcendental Idealism*. New York: Oxford University Press, 1991.
- Weizsacker, Carl Friedrich. "Kant's 'First Analogy of Experience' and Conservation Principle of Physics." *Synthese* 23 (1971), pp. 75–95.
- Weldon, T. D. *Kant's Critique of Pure Reason*. Oxford: Clarendon Press, 1958.
- Werkmeister, William, ed. *Reflections on Kant's Philosophy*. Gainesville: University Presses of Florida, 1975.
- Westphal, Merold. "In Defense of the Thing in Itself." *Kant-Studien* 59 (1968), pp. 118–141.
- Wilkerson, T. E. *Kant's "Critique of Pure Reason": A Commentary for Students*. Oxford: Clarendon Press, 1976.
- Williams, Bernard. "The Certainty of the Cogito," in *Descartes: A Collection of Critical Essays*, ed. by Willis Doney, pp. 88–107. New York: Doubleday, 1967.
- Wilson, Kirk Dalles. "Kant on Intuition." *Philosophical Quarterly* 25 (1975), pp. 247–65.
- Wilson, Margaret. "Leibniz and Materialism." *Canadian Journal of Philosophy* 3 (1974), pp. 495–513.
- Wittgenstein, Ludwig. *Philosophical Investigations*, trans. by G. E. M. Anscombe. Oxford: Basil Blackwell, 1963.
- Wolff, Michael. *Die Vollständigkeit der kantischen Urteilstafel*. Frankfurt am Main: Vittorio Klostermann, 1995.
- Wolff, Robert Paul. *Kant's Theory of Mental Activity*. Cambridge, Mass.: Harvard University Press, 1963.
- Wood, Allen. *Kant's Rational Theology*. Ithaca: Cornell University Press, 1978.
- , ed. *Self and Nature in Kant's Philosophy*. Ithaca: Cornell University Press, 1984.
- Woytowicz, Randolph. *Prolegomena to a Defense of Transcendental Idealism*. Unpublished Ph.D. dissertation, University of California, San Diego, 1994.
- Young, J. Michael. "Construction, Schematism, and Imagination." *Topoi* 3 (1984), pp. 123–31.
- . "Kant's View of Imagination." *Kant-Studien* 79 (1988), pp. 140–64.
- Zocher, Rudolf. "Kants Transzendente Deduktion der Kategorien." *Zeitschrift für philosophische Forschung* 8 (1954), pp. 161–94.

- Zoeller, Guenter. "Making Sense Out of Inner Sense: The Kantian Doctrine as Illuminated by the Leningrad *Reflexion*." *International Philosophical Quarterly* 29 (1989), pp. 263–70.
- . *Theoretische Gegenstandsbeziehung bei Kant: Zur systematischen Bedeutung der Termini "objektive Realität" und "objektive Gültigkeit" in der "Kritik der reinen Vernunft"*. Berlin: de Gruyter, 1984. (*Kant-Studien Ergänzungshefte*, vol. 117.)



## Index

- Abstraction: logical act of, 80, 172, 313
- Actuality, 39–40, 286, 352
- Adickes, Erich, 66
- Aenesidemus, 65
- Affection, 5, 14, 50–51, 64–73, 123, 127, 168; outer, 282–83; self-, 282–84, 290
- Agency, 47–48, 382, 443
- Al-Azm, Sadick, 392
- Alteration, 240–47. *See also* Change
- Ameriks, Karl, 45–47, 335
- Amphiboly of Concepts of Reflection, 59–60, 325, 360, 403
- Analogies of Experience, 15, 34, 40, 225–36, 258, 286; first, 236–46, 250; second, 199, 230, 240, 246–60, 379, 383, 393; third, 260–74; and cognition of objective temporal order, 230–32, 277
- Analytic of Principles, 200, 205
- Anticipations of Perception, 225
- Antinomies of Pure Reason, 22, 109, 321, 332, 443; first, 106, 112, 366–76; second, 321; third, 376–84; fourth, 321; mathematical vs. dynamical, 364–65, 376
- Appearances, 389; inner and outer, 26, 99–108, 124–25; form vs. matter of, 119, 126; temporal relations of, 230–39; transcendental sense of, 285; and phenomena, 54, 57–58; and things in themselves, 21–23, 51–55, 60, 64, 408; totality of, 358–65, 387–89, 397, 408 (*also see* World)
- Appendix, 259, 310, 423–48
- Apperception: the I of, 336–47, 405; transcendental unity of, 61, 87, 163–74, 183–91, 214, 290, 342, 354; synthetic unity of, 135, 164, 168–78, 337–41; analytic unity of, 168, 172; empirical unity of, 183–85, 290; objective unity of, 162, 175–76; analytic principle of, 165–72, 342–43; and inner sense, 277–288

- Apprehension, 198–200, 209, 216, 235, 282, 354; transcendental synthesis of, 79, 170, 185–86, 193–97; empirical synthesis of, 283–84; as successive, 230–34, 250–51, 265
- Aristotle, 85, 120, 133, 137, 155, 240, 244–45, 313
- Attention, 284, 290
- Augustine, 29, 376
- Ausser/in uns*, 24–28, 36, 51–56, 99–101
- Autonomy, 23, 79
- Awareness of self, 277, 289
- Axioms of Intuition, 112, 114, 225
- Backdrop Thesis, 238–39, 291
- Baumgarten, A. G., 120, 262, 333, 427
- Bayle, Pierre, 358, 386
- Beck, J. S., 81–82
- Beck, Lewis White, 90, 219, 246, 255
- Bennett, Jonathan, 38–40, 203, 209, 289, 291, 296, 314, 335, 368, 370–71, 374–75, 380–81
- Berkeley, George, 4–5, 12, 14, 22, 24–27, 35–38, 67, 70–71, 287, 298
- Bird, Graham, 249
- Brandt, Reinhardt, 140
- Buchdahl, Gerd, 116
- Caird, Edward, 238
- Category, 79, 95, 122, 133, 179–80, 185–86, 211–12 *passim*; objective reality of, 217; objective validity of, 220; pure, 17, 56; mathematical vs. dynamical, 199; transcendental vs. empirical use of, 324–27, 335–37, 355; logical and real use of, 159–61, 177–205, 211–14, 220, 335; as conditions of the possibility of experience, 197–201, 220; and logical functions, 133–35, 146–156, 176–77, 266; and schemata, 203–28; and judgment, 179–82, 222–25; and ideas of reason, 316–18
- Causality, 180, 198–99, 261; category and logical function of, 149–51, 271, 363–64; schema of, 223–26, 248–57; and alteration, 246; and necessity, 247, 254–56; principle of, 248, 256–60, 268, 380; reciprocal, 262–73; and Third Antinomy, 376–84; and cosmological proof, 418; and freedom, 377–88; and laws of nature, 378–80, 383, 388; Hume's skepticism regarding, 26–27, 246–47, 252–56, 258; vs. ground-consequent relation, 149–51, 224
- Cause, spontaneity of, 377–78, 382–83
- Change, 237–44, 255–56. *See* Alteration
- Chipman, Lauchlin, 209
- Clarke, Samuel, 102, 373, 380–81, 392
- Clue to the Discovery of all Pure Concepts of the Understanding, *see* Metaphysical Deduction
- Coexistence, 224–25, 231, 241, 260–74
- Cognition: intuitive, vs. discursive, xvi, 28–33, 79; empirical, 424–25; sensible, vs. intellectual, 46, 58, 204; synthetic a priori, 37, 94–96, 116–17, 124, 447; human, conditions of, 24, 35, 37, 39, 68, 147, 159–201, 204; limited to possible experience, 89–93, 95, 307–08, 328, 360–61, 377, 395, 426–27; of self, 275–91; theocentric model of, xv–xvi, 27–38, 44, 73, 79, 89, 121, 127, 132, 373–376, 385, 391, 395, 430, 446; anthropocentric model of, xv–xvi, 34, 37–44, 121; and judgment, 82–83; and reason, 445–48
- Community: mediate vs. immediate, 267–73; category and logical function of, 151, 260–71; schema of, 224–25
- Comparison: logical act of, 80, 172, 313
- Compatibilism, 47–48
- Concept, 78–82 *passim*; mathematical, 204, 208–09, 212, 220; empirical, 79, 92, 95, 103, 204, 208–09, 211, 235, 434–35; pure (*see* Category); problematic, 17; reflected vs. inferred, 313; as analytic unity, 80, 84, 154, 313; form vs. content of, 79–80; objective validity of, 81–85;

- objective reality of, 189; intention vs. extension of, 92, 110; as predicates of possible judgments, 83, 86, 210; as rule, 79–80, 85; exhibition of, 217, 220; and intuition, 78–81, 85, 109–11, 125, 134, 209–11; and schema, 203, 208–09, 221; and judgment, 79, 82–87, 179–80
- Condition: ontological vs. epistemic vs. psychological, xv, 11, 19, 25, 37, 45–47, 174–75, 205, 388; sensible vs. intellectual, 56–57, 159, 203–04, 324; subjective vs. objective, 324, 329; totality of, 315–18, 331, 384–86, 391–93, 430; and unconditioned, 311–321, 327–32, 359–65, 393, 397, 410, 420–21; search for and totality of, as given (*see* Reason)
- Consciousness, 81–83, 160–65; in general, 88, 179–82, 236; empirical, 163–68, 184–85, 193–96, 277, 290–95; transcendental, 167; objective unity of, 87, 175, 182, 349–50; subjective unity of, 87, 162, 175–85, 196
- Contradiction: law of, 90, 397
- Copernican Revolution, xvi, 29, 35–37, 49, 60, 173, 233, 446
- Copernicus, 36
- Cosmology: rational, 320–21, 357–95
- Counterfactuals, 38–41
- Davidson, Donald, anomalous monism, 16
- Dennett, Daniel, 421–422, 439
- Descartes, René, 21–26, 32, 35, 122, 160, 287, 294–95, 298, 333, 344–49, 351–56
- Discursivity Thesis, xiv, 12–16, 27–28, 35, 47, 59, 81, 92, 112, 155, 215, 323–26, 336
- Donnellan, Keith, 374
- Dryer, D. P., 103–104
- Dummett, Michael, 48
- Eberhard, J. A., 30, 71–73, 89–90, 92, 94, 103
- Empiricism, 77, 101–02, 285, 359–60, 382–83, 387
- Epicurus, 359–60, 377, 394
- Essence: nominal vs. real, 31–32
- Event, 247–56, 374–76
- Ewing, A. C., 380
- Existence, 239, 245, 259, 286–87, 293–97, 352; my own, 286–98, 346–56; absolutely necessary, 396–97, 405–19; as a logical predicate, 414
- Experience, 94–95 *passim*; outer and inner, 99–100, 107–08, 266, 278–79, 284–85, 290–97; objective validity of, 197, 248, 285; unity of, 40, 242–44, 382–83, 406–07; a priori principles of, 39–42, 230–36 (*see also* Analogies); conditions of possibility of, 24, 34, 67, 197–201, 236, 242, 252, 257, 261, 267, 274, 406; object of, 24, 252, 267; and perception, 39–40, 179–82, 197–201, 230–35, 251
- Falkenstein, Lorne, 114, 130–32
- Feeling, 81–82, 278
- Focus imaginarius, 425–30, 438–44. *See* Reason
- Freedom, 21, 47–48, 377–78, 381–84, 388; and free will, xix, 377
- Frege, F. L. G., 147
- Function (logical, or Form), 133–56, 147–48, 355; vs. modern concept, 146–47; function vs. form, 147–48; as rule, 148–49; completeness of Table of, 134–46; of quantity, quality, relation, modality, 136–41; moments, 139–42, 145–46; and the categories, 133–35, 146–56, 176–77, 266; of unity in judgment, 137, 141; of judging, 137, 147
- Functionalism, 442
- Garve, Christian, 357
- Geometry, 113, 116–18, 122
- Generation: principle of, 246–47
- Ginsborg, Hannah, 433–34
- God: idea of, 308, 314, 319–21, 327, 358–60, 380, 397; argument for existence of,

God (*continued*)

33–34, 295, 396; ontological proof of existence of, 412–17; cosmological proof of existence of, 417–19. *See also* Ideal

Grier, Michelle, xvii, 330 *passim*

Guyer, Paul, xiv, xix, 7–9, 15, 19, 42, 45–46, 65, 73, 150, 156, 165, 254, 256–57, 269–72, 286, 291, 299–303, 314, 394

Hegel, G. W. F., 422

Heimsoeth, Heinz, 380

Henrich, Dieter, 160–62

Hintikka, Jaakko, 80–81

Hume, David, 11–15, 25–27, 38, 41, 88, 175, 180, 246–47, 252–54, 258, 272, 278, 357, 413, 427–29

Humility: epistemological, xvi, 9, 18, 56, 394

Idea, 313; regulative, 331, 421–22; metaphysical deduction of, 319–22; transcendental deduction of, 437–44; and schema, 438–39. *See also* Reason

Ideal, 332, 398–405, 411, 421; vs. idea, 402–03. *See also* God

Idealism (empirical), 4–5, 21–25, 35, 55, 67, 71, 89, 232, 255, 285–303

Idealism (transcendental), xiv, 4–11, 34–49, 120, 131; proof of, 98, 332, 388–95; as metaphilosophical or formal, xv, 20, 35, 51–52, 61–63, 68, 73, 395, 445; two-aspect view vs. two-world view of, 3, 7, 42–48, 57; and phenomenism, 38–42; vs. transcendental realism, xv, 20–34, 37, 90, 98, 230–33, 285–86, 340–41, 383–88; and resolution of Antinomy, 384–88; and freedom, 383–84

Illusion, 307–08, 321, 349–50; optical, 323–24, 328–32, 340, 426; transcendental, 49, 312, 322–32, 358, 384–85, 404, 420, 435–37; as unavoidable and natural, 308–09, 322, 329–40, 349,

356, 371, 385, 390, 404–06, 410, 419; vs. error, 330–31, 338–40, 404–08, 421, 433; and transcendental realism, xvii, 19, 326, 332–46, 356–58, 385–87, 392–96, 408–09, 420–22, 445–48. *See also* Ideas

Imagination, 185–89, 216; vs. imaging, 187–88, 190; reproductive, 87; transcendental synthesis of, 79, 170–71, 187–95, 218–20, 277–78, 283–84; and figurative synthesis, 186, 189–91, 215, 283; and categories, 191–97; and schemata, 208–15; and time, 189–95; and space, 190–95; and understanding, 152, 186–89; and perception, 187, 210; and Refutation of Idealism, 293–98

Imperatives: categorical, 23, 331, 385

Incongruent counterparts, 118, 130

Induction, *see* Reason hypothetical use

Infinite: series and First Antinomy, 366–71; true transcendental concept of, 366–679; Cantorian conception of, 367–69; defective concept of, 369, 371; space as, 108–13

Intellect: discursive, 166, 169; intuitive, 28–30; pure, 324

Intuition, 64, 78–85, 160, *passim*; sensible or intellectual, 13, 28, 33, 59, 63, 156, 160–67, 354; determinate vs. indeterminate, 81–82, 115, 177; pure, 95–96, 107–12, 122–27, 212; empirical, 96, 107, 123, 177, 193, 212, 352; formal, 112–16, 215; inner, 276; outer, 99–100, 104–06, 117, 119, 124–25, 281–82, 287, 292; form of, 96, 99–100, 112–16, 126; objective validity of, 81; as singular representation, 80–82; as intuited vs. as act of intuiting, 82, 115, 126–27; and concepts, 81–82, 93–95, 109–11, 122, 125, 134, 209; and the categories, 177, 212–13; and judgment, 84–85, 153–55; and the I think, 163–68, 177

Intuitionism: vs. constructivism, 114

I think: the, 336–56

- Jacobi, F. H., 8, 64–68, 73
- James, William, 164
- Judgment (*Urteil*), 79–89, 133–38, 322–23; analytic, 17, 56, 86, 89–95; synthetic, 89, 92–95, 414–15; synthetic a priori, 93–96, 212, 219, 225–28; reflective, 259–60; existential, 414–15; of perception, 179–83, 323–24, 328–29; of experience, 179–83, 212, 225; aesthetic, 81–82; vs. association, 80, 87–88, 175; content vs. logical form of, 90–91, 139; objective validity of, 83–93, 175–83, 200, 248–49, 256; subjective validity of, 178–80; truth of, 205, 248–49, 253, 322–23; and error, 322–23, 329; and logical functions, 135–39, 146–56, 176–78; and categories, 181–82, 222, 224–25; and objects, 173–78, 252; and apperception, 173–82; and Analogies, 225–26; schema judgments, 219–20
- Judgment (power of, *Urteilkraft*), 83, 204–05, 211, 221; vs. ability to understand, 205–06, 210–12; as faculty of subsuming under rules, 205, 206–08, 211–12
- Kant, Immanuel, early philosophy, 32–34, 46, 397, 399–400, 406, 412, 428
- Kant's works: *Anthropology from a Pragmatic Point of View*, 277; *Critique of Judgment*, 82, 216, 221, 259–60, 428, 436; *Critique of Practical Reason*, 23, 26, 48–49, 216; *Groundwork of the Metaphysics of Morals*, 23, 48; Inaugural Dissertation, 46, 57–59, 63, 97–98, 324, 326, 336, 422; *Lectures on Logic*, 78, 80, 83–84, 89, 91, 213, 428; *Metaphysical Foundations of Natural Science*, 176, 237, 244–46, 272, 274; “A New Exposition of the First Principles of Metaphysical Knowledge”, 33; “On a Discovery According to Which Any New Critique of Pure Reason Has Been Made Superfluous by an Earlier One”, 71, 89, 94, 142; “On the Progress of Metaphysics”, 22–23, 400; *Prolegomena to Any Future Metaphysics*, 37, 88, 90–91, 107, 122–24, 178–84, 193, 196, 199–201, 230, 357, 359–60; *Reflexionen*, 86, 89, 278, 279, 286, 290, 294–95, 298–302, 339–40, 400, 428; *Nachlass*, 321; *Träume*, 426; “Lectures on the philosophical doctrine of religion”, 400
- Kemp Smith, Norman, 104, 128, 211, 318, 367, 379–80
- Kneale, William and Martha, 139
- Lambert, J. H., 299
- Langton, Rae, xiv, 7, 9–11, 17–18, 42, 45, 52, 56
- Lavoisier, A.-L., 243
- Leibniz, G. W., 10, 15, 25, 29–32, 43, 71–72, 92–95, 98, 102–03, 106, 120–30, 271–72, 280–81, 324–26, 341, 353–54, 366, 373, 376, 380–81, 392, 398
- Lewis, David, 42
- Locke, John, 15, 21, 30–32, 38, 41, 325
- Logic, 83–85, 140–41, 152–56, 206–07
- Longuenesse, Béatrice, 79, 88, 146–47
- Lovejoy, Arthur, 254
- Maass, J. G., 103, 106
- Malebranche, Nicolas, 29, 32
- Manifold, pure, 113; synthetic unity of, in intuition, 153–54, 161, 173
- Materialism, 22, 24, 342, 352
- Matter, 64, 71–72, 244–46; conservation of, 237, 243–46, 274
- Meier, G. F., 427
- Melnick, Arthur, 113, 268–69, 272
- Mendelssohn, Moses, 299
- Metaphysical Deduction: of categories, 133–34, 140, 152–56, 175, 177, 186, 200, 314–15; of ideas, 319–20
- Metaphysical Exposition, 99, 118, 124–25, 127–28, 131, 276
- Metaphysics, 307, 320–32, 337, 385, 392
- Modality, category of, 363–64, 413–16

- Moral law, 48
- Moore, G. E., 368, 370–71
- Nature, 21, 376–84, 445–48; law of, 185, 245, 256–60, 376–83, 388; causality of, 378–80; systematic unity of, 425, 429, 439–40; vs. world, 364–65, 394
- Necessity: absolute, 396–97, 410–19; and a priority, 235; and causality, 247, 255–56, 270; and essence, 33
- Newton, Isaac, 11, 22, 25–26, 43, 97–98, 102, 120, 123–27, 232, 244, 272–74, 366, 373, 425
- Noumenon, 54, 57–64; and thing in itself, 50, 57–60; causality of, 54, 60, 64–65, 69–73; positive and negative concept of, 52, 58, 63, 324
- Object, 122–23, 173, 233; empirical, 62, 232, 235; transcendental, 50, 59–65, 67–71, 279; of inner sense, 183, 278–80; intelligible, 64–67, 72, 327; of possible experience, 121, 252, 267, 328, 408; in general, 81–82, 120–21, 236, 325–28, 335, 354, 386, 397, 399–400, 406–08; as something in general, 148, 154, 173–75, 279; representation of, 84, 173–75, 233–34; perception of, 162, 251–52; and objectivity, 173–78, 233–34, 248, 251–52
- Ontology, critique of, 324–27
- Paralogisms, 21, 166, 332, 405–08; first, 334–47; second, 341–48; third, 343–468; fourth, 232, 285, 348–51
- Paton, H. J., 86, 126, 211–12, 239, 262, 276, 282–83
- Perception, 167, 186, 189, 197–201, 210, 352; a priori principles of, 40–42; order of, 255–56, 264–66, 271; judgment of, 179–82; of events, 250–56; and categories, 193–201; and synthesis of apprehension, 115, 183, 193–97
- Personal identity, *See* Soul
- Phenomenalism, 4, 12, 16, 25, 38–42, 55, 298
- Phenomenon, 214, 217
- Plato, 313, 359–60, 394
- Possibility: conditions of, 33–34, 398, 406, sum total of, 397–404
- Positivism, 330–32
- Postulates of Empirical Thought, 40, 286
- Practical philosophy, 23, 46–48, 144
- Prauss, Gerold, 50–52, 66, 88
- Prichard, H. A., 6, 17, 202
- Principles of Pure Understanding, 212, 225–26, 309
- Psychology, rational, 279, 285, 308, 320–21, 333–56; empirical, 147, 333
- Quality, category and schema of, 363
- Quantity, category, 203, 362; schema of, 369; of matter, 245
- Realism (Transcendental), 21–34, 55–57, 62, 71, 89, 120–28, 230–33, 326, 332–34, 338–43, 350, 356, 373, 376, 408–09, 420–22 *passim*; as leading to empirical idealism, 21–26
- Realism (Empirical), 121
- Reality, 22, 39, 236, 242, 352; and omnitude realitatis, 399–409; and ens realissimum, 396, 402–20
- Reason, in general, 307–10; as faculty of principles, 309–10, 319; vs. understanding, 307–10, 317, 361–62, 413, 429, 432, 435–36; practical and theoretical, 46–49, 312; ideas of, 312–22, 436–44; principles of, 311–13, 327–30, 413, 424, 431–35; transcendental principle of, 431–32; and logical maxim, 329–31, 339, 404, 430–35, 441, 444; as regulative, 309, 331, 421–48; hypothetical use of (or induction), 425–31, 446–47; transcendent/real use and logical use of, 308, 311–12, 317, 321–22, 327–330, 357, 359, 401; and systematic unity, 309–10, 407–08, 424–48; and rela-

- tional categories, 316–18; and syllogism, 308, 311–15, 318–19, 401–02; euthanasia of, 358, 388, 394, 447; and illusion, 309, 326–32, 392, 418–19, 423–27, 432–33
- Reflection: logical act of, 80, 172, 313; transcendental, 325–28
- Refutation of Idealism, 24, 160, 253, 256, 286–303, 346
- Reid, Thomas, 254
- Reinhold, Karl, 129
- Relation, category, logical function, schemata of, 148–51, 222–25, 280–81
- Russell, Bertrand, 367–70
- Ryle, Gilbert, 150
- Schematism, 190, 202–28; superfluity of, 202–03, 211; task of, 203, 215
- Schema, 17, 34, 208–210, 213–15, 222–25, 236, 242–44; transcendental, 203, 210–28, 251; and categories, 203–28, 222–26, 273, 325; of causality, 248–57; of substance, 236–46; as phenomenon, 214, 217; as formal intuitions, 215–18; as conditions of judgments, 205–06, 213; as transcendental time-determinations, 214–22, 225, 242–43; objective reality of, 242; vs. concept, 203, 208–09; and appearance, 214, 225; and space and time, 214–18, 222–25; and the imagination, 208–09, 214–15
- Shaffer, Jerome, 415
- Shoemaker, Sidney, 344
- Schopenhauer, Arthur, 254, 268, 364, 379–80, 382, 387, 412
- Schultz, J. G., 113
- Science, Kant's philosophy of, 424
- Scotus, Duns, 399
- Self-consciousness, 163–69; empirical, 289; vs. consciousness, 163–64, 176, 288, 344, 349–51; and inner sense, 280; and Refutation of Idealism, 289
- Sellars, Wilfred, 187–188
- Sensation, 127 *passim*. *See* Perception
- Sense (outer), 21–24, 97, 100, 105, 117, 182–83, 190, 194, 277–78, 285, 293–94, 300; ideality of, 280–82; materials of, 280–81. *See also* Space
- Sense (inner), 22–26, 97, 100–01, 105, 167, 185–94, 214–17, 275–85, 300; object of, 278–80, 292–94, 344; vs. apprehension, 277; ideality of, 280–85; materials of, 281. *See also* Time
- Sensibility: human, 161; conditions of, 24–26, 30, 56–57, 119, 214–19; *a priori* forms of, 26–27, 67, 121–31, 189–92; and receptivity, 14, 126, 323; limits of, 58–59, 96; the sensible vs. the intelligible, 30, 52, 69; and the understanding, 191–92, 322–26, 336; and the categories, 160–62, 185; and noumena, 54, 58–60, 63, 325–26
- Separability Thesis, xiv, 6–11, 46
- Simultaneity, *see* Coexistence
- Skepticism, 6, 16, 22–28, 35, 160, 232, 358; Cartesian, 253, 287–98
- Soul (or Self), 278–82, 291, 333–36, 341–43, 346–48, 351; idea of, 308, 314, 319–21, 327, 358–60, 442–43; as it is in itself vs. as it appears, 280–82; hypostatization of, 336–56; and personal identity, 334, 343–45; and persistence, 291–93, 297, 335; and transcendental idealism, 280–85. *Also see* Paralogisms
- Space: as intuition, 108–12, 117–18, 192; as *a priori* intuition, 115–16, 125–27; as formal intuition, 115, 192–94; empirical reality of, 118–20; as form of human sensibility, 21, 99, 121–23, 192–94; as epistemic condition, 24, 98; as form of outer sense, 97, 99–106, 119, 121, 285; a priority of representation of, 99–108, 113–14, 117–18; transcendental ideality of, 21–22, 44, 98, 116–32; as a single unity, 109–10; as infinite, 108–13; substance in, 260–266; as totum analyticum, 369–70; givenness of, 110–16, 192–93; and schemata, 217–18; and

Space (*continued*)

imagination, 190–93; and geometry, 113, 116–18, 122; Newtonian conception of, 25, 97–98, 123–24, 127; Leibnizian conception of, 98, 102–03, 123–24, 127

Spinoza, Benedict de, 16, 27, 29, 105, 151, 321, 353, 403

Spontaneity, 80, 323

Strawson, P. F., xiv, 5–7, 13, 15, 134, 146–147, 150, 156, 168, 187–88, 254–55, 314, 368, 370, 374, 383, 387

Subject: real, 222–23, 242, 245–46; absolute, 334

Substance, 245; category, and logical function of, 148–49, 222–23; schema of, 236–46, 335–37; quantum of, 237, 244–46; and matter, 245–46; states of, 239, 246–62; in space, 260–63; and persistence, 242–46, 299–300; and interaction, 260–62, 266–68; and soul, 334–41, 350

Substratum, 237–246

Subsumption, 96, 205–06, 209, 211–13, 215; syllogistic, 213, 311

Succession, 237–42, 250, 265–70; objective, 248, 253–56, 267–74, 277

Sulzer, J. G., 299

Swedenborg, Emanuel, 426

Synthesis, 163, 168–69 *passim*; intellectual, 171, 386, 391; of apprehension, 79, 115; of imagination, 79, 265; successive, 112–14, 366–70; figurative, 215, 283, 313; consciousness of, 168–76

Synthetic judgment, *see* Judgment

Systematicity: and reason, 309, 424–38; logical principles of affinity, homogeneity, specification, 433–36; and objective validity, 434–38; and transcendental idealism, 445–48

Theology, 308, 320–21, 364, 396–422

Things in themselves, 15, 20–34, 37, 46, 50–57, 64, 67–71, 281, 302, 352, 383–

84, 395; as *considered* in themselves, 16–17, 50–52, 121; relation to noumenon, 50, 57–60; relation to transcendental object, 60–63; as ground or cause of appearances, 54–55, 64–65, 69–73; and time and space, 118–19, 123–25, 127–29, 131

Thing(s) in general, *See* Object: in general

Time, 105–12, 217–18, 222, 238–39, 254, 372–75; as *a priori* intuition, 115–16, 217, 125–26; as formal intuition, 192–94, 216; as form of human sensibility, 21, 99, 115–16, 123, 194; as epistemic condition, 24, 98; as form of inner sense, 97, 167, 185–86, 189, 276–78, 285; as form of all appearances in general, 217, 276–77; a priority of representation of, 99–108, 113–14; transcendental ideality of, 21, 44–45, 69, 98, 300; unity of, 34, 238, 242–44; determination of, 214–17, 226, 236, 243, 277, 299, 302–03; determination of appearances in, 216–17, 232; as totum analyticum, 369–70; as given, 112–16, 192–93; temporal order, 130, 230–36, 237–39, 248–57, 250–52, 264–65, 268, 277; unperceivability of, 216–17, 232, 236–41, 264–66; finite vs. infinite, 112–13, 366–72; in general, 217, 236, 265; and categories, 217–18, 220–28; and imagination, 189–97

Totality, 315–16, 359–61, 366, 370–71

Totum analyticum, 369–70; synthetic, 369–72

Transcendental Aesthetic, 78, 97–98, 120, 185, 192

Transcendental Analytic, 78, 120, 307, 383, 406

Transcendental Deduction, 133, 152, 220, 231, 282, 339; A-edition, 60–62, 71, 79, 159, 164–67, 170, 176, 186, 193, 196, 201, 210; B edition, 83, 87, 114–15, 135, 155, 176, 283–84, 342, 346, 351; B-edition, and compatibility with



- Prolegomena*, 179–82; and two-step proof, 160–62, 184
- Transcendental Dialectic, 78, 307–08, 327–28
- Transcendental Exposition, 99, 116–18, 124, 276
- Truth: and error, 322–29; formal conditions vs. material conditions of, 231, 249; vs. objective validity, 88; empirical, 248–49, 256
- “Two-aspect view” vs. “two-world view”, of transcendental idealism, 3–7, 16
- Understanding, 262 *passim*; as faculty of cognition, 173; as faculty of rules, 205–06, 210–12, 428; pure, 46, 57, 60, 385, 391; logical vs. real or transcendental use of, 152–53; spontaneity of, 14, 80, 176–78, 187, 191, 323; unity of, 407–08; and apperception, 172–73
- Vaihinger, H., 65–67, 73, 129
- Van Cleve, Jay, 42–45, 47, 240
- Verificationism, 373–74
- Walker, Ralph, 446–47
- Walsh, W. H., 82, 330–31, 314, 391–93, 430
- Warnock, J. G., 203, 208
- Watkins, Eric, 262, 272–73
- Whole, 263, 271, 369–70, 389–91, 395; mathematical vs. dynamical, 364–65, 377–81, 393
- Williams, Bernard, 353
- Wittgenstein, Ludwig, 73, 187, 206
- Wolff, Christian, 120, 333
- Wolff, Michael, 140, 143
- Wood, Allen, 399–400, 402, 415
- World, 287, 308, 314, 319–20, 327, 366, 370, 387–95, 409; infinite or infinite, 362, 366–76, 386–89; vs. nature, 364, 394; idea of, 358–61, 443; and categories, 362–64; as a totum syntheticum, 370–72, 394
- Young, Michael, 188–90
- Zeno, the Eleatic, 386–87

# Kant's Transcendental Idealism

An Interpretation and Defense REVISED AND ENLARGED EDITION

Henry E. Allison

This landmark study is now reissued in a new edition that has been vastly rewritten and updated to respond to recent Kantian literature. It includes a new discussion of the Third Analogy, a greatly expanded discussion of Kant's Paralogisms, and entirely new chapters dealing with Kant's theory of reason, his treatment of theology, and the important Appendix to the Dialectic.

"Henry Allison's *Kant's Transcendental Idealism* is an influential and groundbreaking book that provoked both textual and philosophical criticisms. This second edition amounts to Allison's detailed, impressive answer to these charges, and the result is almost a new book altogether, as thoughtful, original, and important as the first edition. In addition, Allison's revisions and new discussions of such topics as the discursivity of human knowledge, the nature of human reason, and the argument of the Third Analogy all venture into wholly new territory, where Allison again succeeds in effectively reframing the basic issues at stake in very helpful, novel ways, while remaining faithful to both the spirit and the letter of the Kantian text." ROBERT PIPPIN, University of Chicago

"This second edition is a very significant expansion of the original classic, with updated discussions of the latest scholarship and a much-enlarged treatment of the Transcendental Dialectic. A must buy." KARL AMERIKS, author of *Interpreting Kant's Critiques*

Praise for the earlier edition:

"Probably the most comprehensive and substantial study of the *Critique of Pure Reason* written by any American philosopher. . . . This is a splendid book."  
LEWIS WHITE BECK

"A sensitive, well-organized, philosophically ingenious, and indispensable work."  
ARTHUR MELNICK, *Philosophical Review*

Henry E. Allison is professor of philosophy at Boston University.

Yale University Press New Haven & London

yalebooks.com

ISBN 0-300-10266-6



9 780300 102666